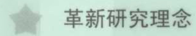


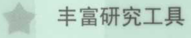
格致方法·定量研究系列　吴晓刚　主编

世代分析（第二版）

[美] 诺瓦尔·D. 格伦（Norval D. Glenn）著

於嘉　译


革新研究理念


丰富研究工具

最权威、最前沿的定量研究方法指南

25

格致出版社　上海人民出版社



格致方法·定量研究系列

1. 社会统计的数学基础
2. 理解回归假设
3. 虚拟变量回归
4. 多元回归中的交互作用
5. 回归诊断简介
6. 现代稳健回归方法
7. 固定效应回归模型
8. 用面板数据做因果分析
9. 多层次模型
10. 分位数回归模型
11. 空间回归模型
12. 删截、选择性样本及截断数据的回归模型
13. 应用logistic回归分析（第二版）
14. logit与probit：次序模型和多类别模型
15. 定序因变量的logistic回归模型
16. 对数线性模型
17. 流动表分析
18. 关联模型
19. 中介作用分析
20. 因子分析：统计方法与应用问题
21. 非递归因果模型
22. 评估不平等
23. 分析复杂调查数据（第二版）
24. 分析重复调查数据
25. 世代分析（第二版）
26. 纵贯研究（第二版）
27. 多元时间序列模型
28. 潜变量增长曲线模型
29. 缺失数据
30. 社会网络分析（第二版）
31. 广义线性模型导论
32. 基于行动者的模型
33. 基于布尔代数的比较法导论
34. 微分方程：一种建模方法
35. 模糊集合理论在社会科学中的应用
36. 图形代数
37. 项目功能差异

上架建议：社会研究方
ISBN 978-7-5432-2175



定价：15.00元
易文网：www.ewen.cc
格致网：www.hibooks.cn

格致方法·定量研究系列　吴晓刚　主编

世代分析（第二版）

[美] 诺瓦尔·D.格伦(Norval D.Glenn) 著
於　嘉 译

SAGE Publications ,Inc.

格致出版社　上海人民出版社

图书在版编目(CIP)数据

世代分析:第2版/(美)格伦(Glenn, N. D.)著;於嘉译. —上海:格致出版社:上海人民出版社,2012

(格致方法·定量研究系列)

ISBN 978-7-5432-2175-8

Ⅰ.①世… Ⅱ.①格… ②於… Ⅲ.①社会科学-研究方法 Ⅳ.①C3

中国版本图书馆CIP数据核字(2012)第219737号

责任编辑　罗　康

格致方法·定量研究系列

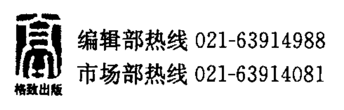
世代分析(第二版)

[美]诺瓦尔·D.格伦 著

於嘉 译

出　版　世纪出版集团 www.ewen.cc　格致出版社 www.hibooks.cn　上海人民出版社

(200001 上海福建中路193号24层)



编辑部热线 021-63914988

市场部热线 021-63914081

发　行　世纪出版集团发行中心

印　刷　浙江临安曙光印务有限公司

开　本　920×1168毫米 1/32

印　张　3.25

字　数　59,000

版　次　2012年10月第1版

印　次　2012年10月第1次印刷

ISBN 978-7-5432-2175-8/C·91

定　价　15.00元

出版说明

由香港科技大学社会科学部吴晓刚教授主编的"格致方法·定量研究系列"丛书，精选了世界著名的 SAGE 出版社定量社会科学研究丛书中的 35 种，翻译成中文，集结成八册，于 2011 年出版。这八册书分别是：《线性回归分析基础》、《高级回归分析》、《广义线性模型》、《纵贯数据分析》、《因果关系模型》、《社会科学中的数理基础及应用》、《数据分析方法五种》和《列表数据分析》。这套丛书自出版以来，受到广大读者特别是年轻一代社会科学工作者的欢迎，他们针对丛书的内容和翻译都提出了很多中肯的建议。我们对此表示衷心的感谢。

基于读者的热烈反馈，同时也为了向广大读者提供更多的方便和选择，我们将该丛书以单行本的形式再次出版发行。在此过程中，主编和译者对已出版的书做了必要的修订和校正，还新增加了两个品种。此外，曾东林、许多多、范新光、李忠路协助主编参加了校订。今后我们将继续与 SAGE 出版社合作，陆续推出新的品种。我们希望本丛书单行本的出版能为推动国内社会科学定量研究的教学和研究作出一点贡献。

总序

往事如烟，光阴如梭。转眼间，出国已然十年有余。1996年赴美留学，最初选择的主攻方向是比较历史社会学，研究的兴趣是中国的制度变迁问题。以我以前在国内所受的学术训练，基本是看不上定量研究的。一方面，我们倾向于研究大问题，不喜欢纠缠于细枝末节。国内一位老师的话给我的印象很深，大致是说：如果你看到一堵墙就要倒了，还用得着纠缠于那堵墙的倾斜角度究竟是几度吗？所以，很多研究都是大而化之，只要说得通即可。另一方面，国内（十年前）的统计教学，总的来说与社会研究中的实际问题是相脱节的。结果是，很多原先对定量研究感兴趣的学生在学完统计之后，依旧无从下手，逐渐失去了对定量研究的兴趣。

我所就读的美国加州大学洛杉矶分校社会学系，在定量研究方面有着系统的博士训练课程。不论研究兴趣是定量还是定性的，所有的研究生第一年的头两个学期必须修两门中级统计课，最后一个学期的系列课程则是简单介绍线性回归以外的其他统计方法，是选修课。希望进一步学习定量研

究方法的可以在第二年修读另外一个三学期的系列课程，其中头两门课叫“调查数据分析”，第三门叫“研究设计”。除此以外，还有如“定类数据分析”、“人口学方法与技术”、“事件史分析”、“多层线性模型”等专门课程供学生选修。该学校的统计系、心理系、教育系、经济系也有一批蜚声国际的学者，提供不同的、更加专业化的课程供学生选修。2001 年完成博士学业之后，我又受安德鲁·梅隆基金会资助，在世界定量社会科学研究的重镇密歇根大学从事两年的博士后研究，其间旁听谢宇教授为博士生讲授的统计课程，并参与该校社会研究院(Institute for Social Research)定量社会研究方法项目的一些讨论会，受益良多。

2003 年，我赴港工作，在香港科技大学社会科学部，教授研究生的两门核心定量方法课程。香港科技大学社会科学部自创建以来，非常重视社会科学研究方法论的训练。我开设的第一门课“社会科学里的统计学”(Statistics for Social Science)为所有研究型硕士生和博士生的必修课，而第二门课“社会科学中的定量分析”为博士生的必修课(事实上，大部分硕士生在修完第一门课后都会继续选修第二门课)。我在讲授这两门课的时候，根据社会科学研究生的数理基础比较薄弱的特点，尽量避免复杂的数学公式推导，而用具体的例子，结合语言和图形，帮助学生理解统计的基本概念和模型。课程的重点放在如何应用定量分析模型研究社会实际问题上，即社会研究者主要为定量统计方法的“消费者”而非“生产者”。作为“消费者”，学完这些课程后，我们一方面能够读懂、欣赏和评价别人在同行评议的刊物上发表的定量研究的文章；另一方面，也能在自己的研究中运用这些成熟的

方法论技术。

上述两门课的内容，尽管在线性回归模型的内容上有少量重复，但各有侧重。“社会科学里的统计学”（Statistics for Social Science）从介绍最基本的社会研究方法论和统计学原理开始，到多元线性回归模型结束，内容涵盖了描述性统计的基本方法、统计推论的原理、假设检验、列联表分析、方差和协方差分析、简单线性回归模型、多元线性回归模型，以及线性回归模型的假设和模型诊断。“社会科学中的定量分析”则介绍在经典线性回归模型的假设不成立的情况下的一些模型和方法，将重点放在因变量为定类数据的分析模型上，包括两分类的 logistic 回归模型、多分类 logistic 回归模型、定序 logistic 回归模型、条件 logistic 回归模型、多维列联表的对数线性和对数乘积模型、有关删节数据的模型、纵贯数据的分析模型，包括追踪研究和事件史的分析方法。这些模型在社会科学研究中有着更加广泛的应用。

修读过这些课程的香港科技大学的研究生，一直鼓励和支持我将两门课的讲稿结集出版，并帮助我将原来的英文课程讲稿译成了中文。但是，由于种种原因，这两本书拖了四年多还没有完成。世界著名的出版社 SAGE 的“定量社会科学研究”丛书闻名遐迩，每本书都写得通俗易懂。中山大学马骏教授向格致出版社何元龙社长推荐了这套书，当格致出版社向我提出从这套丛书中精选一批翻译，以飨中文读者时，我非常支持这个想法，因为这从某种程度上弥补了我的教科书未能出版的遗憾。

翻译是一件吃力不讨好的事。不但要有对中英文两种

语言的精准把握能力,还要有对实质内容有较深的理解能力,而这套丛书涵盖的又恰恰是社会科学中技术性非常强的内容,只有语言能力是远远不能胜任的。在短短的一年时间里,我们组织了来自中国内地及港台地区的二十几位研究生参与了这项工程,他们目前大部分是香港科技大学的硕士和博士研究生,受过严格的社会科学统计方法的训练,也有来自美国等地对定量研究感兴趣的博士研究生。他们是:

香港科技大学社会科学部博士研究生蒋勤、李骏、盛智明、叶华、张卓妮、郑冰岛,硕士研究生贺光烨、李兰、林毓玲、肖东亮、辛济云、於嘉、余珊珊,应用社会经济研究中心研究员李俊秀;香港大学教育学院博士研究生洪岩璧;北京大学社会学系博士研究生李丁、赵亮员;中国人民大学人口学系讲师巫锡炜;中国台湾"中央"研究院社会学所助理研究员林宗弘;南京师范大学心理学系副教授陈陈;美国北卡罗来纳大学教堂山分校社会学系博士候选人姜念涛;美国加州大学洛杉矶分校社会学系博士研究生宋曦。

关于每一位译者的学术背景,书中相关部分都有简单的介绍。尽管每本书因本身内容和译者的行文风格有所差异,校对也未免挂一漏万,术语的标准译法方面还有很大的改进空间,但所有的参与者都做了最大的努力,在繁忙的学习和研究之余,在不到一年的时间内,完成了三十五本书、超过百万字的翻译任务。李骏、叶华、张卓妮、贺光烨、宋曦、於嘉、郑冰岛和林宗弘除了承担自己的翻译任务之外,还在初稿校对方面付出了大量的劳动。香港科技大学霍英东南沙研究院的工作人员曾东林,协助我通读了全稿,在此

我也致以诚挚的谢意。有些作者，如香港科技大学黄善国教授、美国约翰·霍普金斯大学郝令昕教授，也参与了审校工作。

我们希望本丛书的出版，能为建设国内社会科学定量研究的扎实学风作出一点贡献。

吴晓刚

于香港九龙清水湾

序

许多社会科学学科都以“世代”研究为教学重心，例如人口统计学、流行病学、心理学、政治科学和社会学。格伦·埃尔德(Glen Elder)曾在其著作《大萧条中的孩子们——社会变迁对人生的影响》(*Children of Great Depression——Social Change in Life Experience*)一书中对“世代”这一概念进行了深入研究。该书讲述了一群孩子在成长过程中种种坎坷的经历，对他们而言，现代社会许多司空见惯的东西都是遥不可及、需要拼命争取才能拥有的。在本书的讨论中，我们以“迷惘的一代”为例，讲述他们怎样在战后努力向社会证明自己的能力。正是这些独特的经历造就了关于世代的研究。

一般而言，世代分析属于定量研究的范畴，研究者利用很多世代概念的估测方法对一些结果变量的效应进行研究。例如，在大萧条时期，经历过战争的社会群体身上的某些行为特征。从这种意义上来讲，世代的成员都必须在特定时间段有同样的经历。另一方面，同一年出生的人(或特定时期出生的人群)经历了同样的社会变迁，或多或少都会表现出

相似的性格特征。不同年龄的人有时也表现出这种群体特征。例如,同一年开始实习的毕业生们。所以在大多数情况下,“世代”指的是某一特定的出生人群。

20 世纪 70 年代出版的《大萧条中的孩子们——社会变迁对人生的影响》点燃了研究者的研究热情,他们开始寻找估测特定社会行为或社会现象以及估计世代效应的方法。可这并不简单,因为估测对象不仅仅包括世代(出生年份)的效应,还有年龄(出生后)以及特定时期(现有年份)的效应。众所周知,如果无法识别限制,那么 APC 线性模型和叠加模型也就无法被识别,因为年龄、时期和世代三个因素恰好构成一个线性函数式,其中年龄＝时期－出生年,为一个恒等式。

诺瓦尔·D. 格伦发表的《世代分析》一书中提供了几种应对识别问题的新方法。作者在书中对世代分析进行了更深层次的讨论,内容涉及方法论基础——包括数据问题——并详细讲述了几种可供研究者选择的 APC 估测法。社会科学各个学科的研究者都会在此书中获得很多收获,那些研究年龄变化以及社会文化变迁的学者们将是最大的受益人。

廖福挺

目 录

第1章

引　言

世代分析并不是一项统计技术,而是检验数据时经常采用的策略。在过去的几十年中,随着可用数据量的不断增加,越来越多的社会科学研究者开始关注世代分析。虽然世代分析可以有效地应用于研究与年龄增长及社会文化变革相关的重大问题,但显然更多研究者是被其异常神秘的方法论难题所吸引的。毫无疑问,这种兴趣并非完全有益。许多针对世代分析的研究只是练兵,研究者花了许多精力去证明的问题其实在逻辑上不可行。这样的研究导致了许多过于自信,却很可能不正确的结论。

因此,我写这本书的主要目的,一方面是阻止研究者怀着这种错误的观念进行机械化的世代研究,另一方面是为了鼓励正确的研究方式。能更深入地了解世代分析的方法自然很好,但前提是研究者必须谨慎,且拥有足够的专业经验,这样才能正确破解其中的奥妙。

第1节 世代研究的目的

一般来说，研究者会在两种情况下使用世代分析，虽然这两种情况都有非常明显的不同目的，但还是引发不少人对研究的目的进行一些区分。

在多数情况下，世代研究主要用于估测人类年龄变化所产生的效应，有时也估测婚姻产生的效应。显然，随机实验最适合用来估测诱因，但不能用于研究年龄变化产生的效应。我们无法将一定数量的个体随机分成实验组和对照组，无法观测实验组个体年龄变化的过程，并保证对照组不受与年龄变化相关的刺激因素的影响。因此，研究年龄变化效应的学者只能在一些非实验方法和准实验方法中进行选择，但这些方法都有一定的局限性，世代分析有独特的优点。我将在本章的下一节中讨论几种主要的方法。而本书的主要侧重点在于讨论与年龄变化相关的研究，但文中也经常出现将世代用于分析其他情况的讨论。

世代分析的第二大用途是帮助研究者理解社会、文化及政治变革的根源和性质。这也是为什么世代分析在近几年中迅速成为关注焦点的原因之一。也许将来，这种用途会比前一种更加普遍。当然，除了世代法，肯定还有其他分析方法能够在社会和文化变迁的定量研究中发挥作用，但是世代

研究所能达到的深度绝对是独一无二的,在这方面,没有任何一种方法能与之媲美。因此,我不会在本书中将世代研究和其他方法进行系统化的比较。尽管本书着重于阐释世代分析在研究年龄变化中的作用,但是我仍然会用一章的篇幅来讨论世代方法是怎样用于研究社会变迁的。

第2节 | 定义:世代研究与相关研究方法之间的比较

英文单词“cohort”原本用来形容士兵队伍,现在有时用来泛指一群拥有某些相同特质的人。而在本书及其他同期群分析的文献中,这个词汇被用来形容在特定时间段中经历过特定事件的人群。“世代”是社会学研究中最常见的同期群研究对象,所谓“世代”,就是指特定年份、某个10年间或其他时间段内出生的人群。如果没有限定词特指,那么文献中所说的同期群研究就是世代研究。在其他情况下,研究者用特定事件来特指同期群研究,比方说,曾经有研究者以退休组和毕业组为对象。如果同期群不是人类,那就必须增加一个限定词。同期群的种类很多,有以婚姻,也有以参加组织活动来定义同期群。虽然同期群的指代对象一般都是世代[①],但也可以是婚姻、组织、课本、电影、车模,或其他任何在特定时间段内发生的事件。

很不幸的是,社会科学文献中居然曾出现“年龄世代”这样的定义,由于缺少事件因素,这种说法是不符合规则的。

---

① “cohort analysis”有多种翻译,如队列分析、出生组分析、同期群分析、同辈群体分析、世代分析等。因为后文涉及的主要分析对象都是出生世代,而研究中最常见的应用也是出生世代,所以本节中统 译为“世代分析”。 译者注

年龄当然是条件之一,而且是不断变化的条件,但是年龄绝对不是时间。如果研究中出现“年龄世代”这样的词汇,我们就很难确定其真正的含义——是指世代,还是简单的表示某一年龄段的人?

这两种含义是截然不同的。打个比方,所有出生在20世纪60年代的人被称为“60世代”(“60后”),但是根据研究发生的时间,所有个体当时的年龄都是不同的。而只有30岁的人才能以年龄归为一类。也就是说,世代每往前退一年,年龄组都会发生变化。

在通常情况下,考虑到一个因变量需要两个时间点的信息才能被测量,所以只有两个或两个以上的世代进行对比的研究才能称为“世代分析”。不是所有以某一世代为对象的研究都是世代研究。比如从2002年11月开始,有研究者分别对青年群体、中年群体和老年群体的政治态度进行调查,并进行对比;也有人在1995年将某高校1985年所有毕业生的政治态度进行对比。这两项研究都不是真正的世代研究。前一项是横截面研究,数据从一个时间点或时间段中收集,后者是固定样本跟踪研究,为了测量同一群体的某些特征而在多个时间点收集数据。两种研究各有所长,都是非常有价值的研究,但两者都无法满足绝大多数世代研究的要求,无法测量年龄增长所带来的影响,也就是年龄效应。

横截面数据也能显示年龄所导致的差别,但不能反映年龄效应,因为不同年龄的人(或其他研究对象)本身就属于不同的世代,其经历不同,产生的影响也不同。换句话说,这些不同就是所谓的“世代效应”。举例来说,20世纪20年代出生的美国人在青少年时期经历了大萧条,而10年之后出生

的人在他们的青少年时期见证了美国前所未有的经济增长。我们有理由相信,这种青少年经历的不同会导致两个世代在经济政治态度与行为上有着终生的差异。在出生率较低年代出生的人被称为“小型生育世代”,他们一生中的经济生活、教育、事业和婚恋的机遇和大型生育世代的情况存在很大区别,而这些区别也影响着他们各自的态度和行为。因此,之所以不能用横截面研究来测量年龄效应或世代效应,是因为它会混淆研究结果。比方说,研究者无法弄清导致中年人和老年人之间观念区别的真正原因,无法判断到底是因为老年人年长,其经历有别于中年人,还是因为各自的出生年代不同而导致的差异?

横截面研究不适合在年长和年轻的存活个体间测量年龄效应的另一个原因在于因变量与寿命之间的相关性。例如,美国21世纪初的全国调查显示,中年人饮用酒精饮料的比例高于老年人。这个发现本身并不能证明随着男性年龄增长,酒精饮料的饮用量会减少。这种差别当然可能部分是由世代效应导致,但也可能是一种组成效应,因为饮酒较多的人可能比较少饮酒或从不饮酒的人活得更短。大多数世代分析都使用从一系列截面研究中获得的数据,因此他们将不同死亡率的效应与年龄或世代效应混淆。

固定样本跟踪研究可以验证由死亡率导致不同的组成效应,同时它们也可以处理不断变老的世代中的个体层面和汇总层面的变化。但是,固定样本跟踪研究自身也不能为年龄效应提供具有说服力的证据,因为并不是所有随着个体年龄增长而出现的变化都是由于个体老龄化所导致的。尤其是在现代社会中,人们并不是在一个静止的社会中不断变

老,而是处在一个不断变化的社会中。随着个体年龄的增长,社会或文化变化的影响也作用于他们,这便造成了态度、行为、健康和情感的变化,并抵消了年龄增长在静止社会中的效应。在跟踪数据研究中,这些时期效应与年龄效应相混淆。例如在美国,在20世纪70年代进入青年期的世代可能在20世纪80年代后期变得相对保守。我们无法判断这样的转变有多少是源自老龄化,因为整个社会也在向着同样的方向变化,那么就不能分辨世代间的趋势是不是源于时期的影响。因此,对20世纪70年代中期及10年后的高中生的政治态度的跟踪研究并不能证明是否有因转变为成年人而带来的影响。

用固定样本跟踪来研究年龄效应的另一个局限在于,研究群体可能受他们在调查中的参与情况的影响。如果这种效应发生在因变量上,那么这就是一个固定样本跟踪研究的调适效应。在跟踪数据中,这一效应与年龄和时期效应会混淆。另一种阐释这个问题的方式,即如果存在固定样本跟踪研究的调适效应,随着研究和调查(重复收集数据)次数的增加,一个原本能代表整体的样本将变得不再那么具有代表性。

固定样本跟踪研究的条件习惯效应可能发生在很多情况下。例如,如果因变量是根据一个标准测验测量出的能力或者技能,则可能存在一种特殊的固定样本条件习惯效应,称为“练习效应”,也就是说,随着固定样本跟踪研究一次又一次的进行,研究对象可能适应了测试而有更好的表现。同样,不断重复询问有关某一主题的尖锐问题,可能使受访者进行深思从而导致态度上的变化。或者,向受访者详细询问其工作、婚姻或者其他生活方面,可能会使他们渐渐对这类问题产生抗拒。受研究对象年龄增长影响的固定样本跟踪

调查非常容易受固定样本条件习惯效应的影响,因为受访者可能习得了有效的方法来对他们在研究中的行为进行调节。

我们可以对固定样本追踪数据进行世代分析,但是很少有研究这样做,而且使用两个或两个以上截面研究的一般的世代分析可以避免固定样本调适效应。在这种称做"重复截面"的设计下,没有个体在多于一个时点接受了调查。相反,研究对来自每个世代的不同个体组成的样本在不同的时间进行了研究。

解释世代分析逻辑的一个启示性工具就是标准世代表,它通过将不同截面数据并排放置,显示出年龄与其他因变量之间的关系,而年龄的区间与进行数据调查的区间是一样的。表 1.1 正显示了一个这样的例子,它显示了美国人口普查数据中已婚妇女的比例。

表 1.1 根据年龄和年份分类的女性已婚的比例(美国)

年 龄	年 份			
	1968	1978	1988	1998
25—34	87.4	76.6	67.3	67.3
35—44	87.1	82.1	76.3	72.1
45—54	82.4	80.5	76.2	70.8
55—64	67.7	70.4	70.7	67.8
65—74	46.5	48.3	53.3	54.8

资料来源:美国人口普查局进行的长期人口调查。百分比是根据美国统计局提供的数据计算得到(1969:表 37; 1979:表 51; 1990:表 49; 1999:表 63)。

这个表中的每一列都是一个截面数据,而其中年龄、世代和一部分组成效应相互混淆。每一行都是处于同一年龄的四个不同的世代,其中,时期和世代效应相互混淆。除了 1998 年年龄在 25—34 岁之间和 1968 年年龄在 65—74 岁之

间的世代,表中的每一个世代都可以跟踪至少 10 年,即表中从左上方至右下方对角线所显示的。在每一个世代对角线数据中,年龄和时期效应都相互混淆,而这个例子中的一小部分组成效应可以在中年后进行追踪。

与来自截面研究的数据一样,标准世代表中的每一列都受到效应混淆的影响,而每一个世代对角线的数据都与来自固定样本追踪研究的数据一样,会将年龄和时期混淆。然而,在世代表中,有许多列和许多世代对角线,这使得研究者希望能够通过统计方法分离出年龄、时期和世代效应。多年前,我将试图利用统计方法分离出世代数据中各种效应的方法称做“徒劳的追寻”(Glenn, 1976),除非在极少数的情况下所有的效应都是非线性的。然而,并不是所有人都放弃了这种尝试。试图寻找一个能够一劳永逸地应用于估计效应的统计方法,是科学史上重复进行那些逻辑错误的尝试中最讽刺的例子之一。

第3节 识别问题

由于不可能利用统计方法分离出年龄、时期和世代效应，这就产生了识别问题，这一问题发生在当有3个或3个以上的自变量需要纳入研究，而每一个都是其他的线性函数的时候。换句话说，每一个自变量与其他自变量的复相关是唯一的——这也是共线性中可能出现的最极端的例子。如果所有的自变量都作为预测因子纳入回归方程或者类似的分析，则电脑程序不能运行。在其他变量都被控制的情况下，每个剩余自变量的方差均为0。例如，当我们有理由认为两个变量间的差别和两个变量本身都将影响因变量时，识别问题便发生了。一个例子即对婚姻质量的研究，丈夫的特征、妻子的特征和夫妻之间的差异都可能影响婚姻质量。另一个例子是社会流动性对幸福感的心理效应的研究，其中，出身的阶层、最终归属的阶层和两者的差别都可能影响到幸福感。

在世代分析中，三个相互关联的变量当然是年龄、时期和世代，其中每一个都是其他两个的完全线性函数。也就是说，根据其中两个变量的信息就可以推知第三个变量的信息。例如，如果我们知道一个在1990年进行访谈的调查对象当时是20岁，我们就会知道她的出生年月以及她的世代成员信息。或者，如果我们知道一个1990年进行访问的调

查对象出生于 1970 年,则我们可以推知这个受访者当时的年龄为 20 岁。所以,如果时期和世代(或其他 3 个中的任意两个变量)都已经作为预测变量被纳入回归方程,那么再加入年龄就会造成冗余信息,使得电脑程序无法运行。如果有理由相信年龄、时期或者世代的任意一个没有效应,则剩余两个变量的效应很容易被估计,但是并没有直接的方法可以同时估计出这三个变量的效应。

识别问题的本质可以由表 1.2、表 1.3 和表 1.4 中的假设性数据所解释,其中对每个表格中因变量值的线性变异最简单的解释是,它们各自仅表现了年龄效应(表 1.2)、时期效应(表 1.3)和世代效应(表 1.4)(为了便于解释,我们假设因变量与寿命无关,因此也没有由不同的死亡率导致的组成效应)。然而,每个表中的数据同样适用于两个变量的解释,正如每个表底部的其他解释所显示的那样,而每个表中的变异模式都可能来源于年龄、时期和世代效应之间无限的组合。显而易见,我们无法利用统计方法本身从这些不同的效应组合中选择可以产生相同数据的组合。研究者的这种选择必须基于理论或者他们对现象来源的认识,而非基于数据本身。换句话说,选择必须基于菲利普·康威提出的"附加信息"(Philip Converse, 1976)。这种信息有时来自数据中记录的附加变量(除了年龄、时期、世代和因变量以外的变量),但更多时候来源于其他途径。

研究者可能认为,尽管在逻辑上可以用 1 至 3 个变量对于世代数据线性模式的变异进行解释,但一个变量的解释往往是最有效的。然而事实并非如此。即便当变异的模式像表 1.2、表 1.3 和表 1.4 那么简单,2 个或 3 个变量的解释往

往与理论或者其他来源的数据更加契合。此外，哪怕只是简单的线性变异，也往往比表格中的假设数据复杂很多，而复杂的模式就需要复杂的解释。

表 1.2 单纯显示年龄效应的数据模式(抵消了时期和世代效应，或者年龄效应与相互抵消的时期和世代效应的混合)

年 龄	年份					
	1950	1960	1970	1980	1990	2000
20—29	50	50	50	50	50	50
30—39	55	55	55	55	55	55
40—49	60	60	60	60	60	60
50—59	65	65	65	65	65	65
60—69	70	70	70	70	70	70
70—79	75	75	75	75	75	75

注：单元格中的数据是某个因变量的假设值。几个解释为：(1)每 10 年的衰老将导致因变量 5 点的增加；(2)每 10 年的变化对因变量产生 5 点的正向时期效应和 5 点的负向世代效应；(3)每 10 年的变化使因变量存在一部分年龄效应和相互抵消的时期和世代效应的混合。这些效应的有限组合可以生成表中所示的因变量的变异模式。

表 1.3 单纯显示时期效应的数据模式(抵消了年龄和世代效应，或者时期效应与相互抵消的年龄和世代效应的混合)

年 龄	年份					
	1950	1960	1970	1980	1990	2000
20—29	30	35	40	45	50	55
30—39	30	35	40	45	50	55
40—49	30	35	40	45	50	55
50—59	30	35	40	45	50	55
60—69	30	35	40	45	50	55
70—79	30	35	40	45	50	55

注：单元格中的数据是某个因变量的假设值。几个解释为：(1)每 10 年的变化将导致因变量 5 点的增加；(2)每 10 年的时间变化对因变量有 5 点的正向年龄效应和 5 点的正向世代效应；(3)每 10 年的变化使因变量存在一部分时期效应和互相抵消的年龄和世代效应的混合。这些效应的有限组合可以生成表中所示的因变量的变异模式。

表 1.4 单纯显示世代效应的数据模式(抵消了年龄和时期效应，或者世代效应与相互抵消的年龄和时期效应的混合)

年龄	年份					
	1950	1960	1970	1980	1990	2000
20—29	50	55	60	65	70	75
30—39	45	50	55	60	65	70
40—49	40	45	50	55	60	65
50—59	35	40	45	50	55	60
60—69	30	35	40	45	50	55
70—79	25	30	35	40	45	50

注:单元格中的数据是某个因变量的假设值。几个解释为:(1)每一个到达了成年的世代将比之前的世代的因变量高出 5 点;(2)每 10 年的变化对因变量有 5 点的负向年龄效应和 5 点的正向时期效应;(3)每 10 年的变化使因变量存在一部分世代效应和互相抵消的年龄和时期效应的混合。这些效应的有限组合可以生成表中所示的因变量的变异模式。

例如,美国的年龄与工作满意度这个案例(见表 1.5)。工作满意度的整体水平在过去的几十年里相当稳定,且年龄与工作满意度之间有稳定的正向相关关系。对于这个稳定关系的最简单解释就是,它反映了年龄效应或者年龄效应与由于进出劳动力市场导致的组成效应的混合。

表 1.5 表示对工作“非常满意”[a] 的人数百分比[b]

年份	年龄					
	25—34	35—44	45—54	55—64	65—74	总计
1972—1976	44.4 (1475)	53.9 (1208)	51.8 (1358)	60.1 (936)	63.4 (313)	52.4 (5290)
1977—1980	41.3 (989)	49.0 (769)	52.2 (684)	61.1 (556)	60.6 (223)	50.2 (3221)
1982—1986	42.4 (1868)	49.1 (1384)	54.1 (1106)	57.2 (865)	62.4 (320)	52.9 (5543)
1987—1991	40.1 (1614)	46.5 (1553)	50.9 (1062)	52.2 (655)	57.5 (277)	49.4 (5161)
1993—1996	39.3 (1483)	44.6 (1563)	47.3 (1246)	55.9 (584)	68.1 (183)	45.9 (5059)

续表

年份	年龄					
	25—34	35—44	45—54	55—64	65—74	总计
1998—2002	46.3 (1275)	45.6 (1394)	49.1 (1149)	56.5 (539)	61.6 (189)	48.7 (4546)
总计	42.2 (8704)	47.8 (7871)	50.8 (6605)	57.3 (4135)	62.1 (1505)	48.9 (28820)

注：a. 其他回答为"一般满意"、"不太满意"和"非常满意"。
b. 通过加权近似获得了受访者的人数。权重是家庭中成年人的比例除以GSS家庭中的成年人平均数量。
资料来源：全国民意研究中心(NORC)于1972—2002年进行的年度美国综合社会调查(GSS)。

表1.6 因变量的非线性变异可以解释为仅仅反映了年龄或者组成效应的数据模式

年龄	年份					
	1950	1960	1970	1980	1990	2000
20—29	50	50	50	50	50	50
30—39	52	52	52	52	52	52
40—49	62	62	62	62	62	62
50—59	62	62	62	62	62	62
60—69	50	50	50	50	50	50
70—79	45	45	45	45	45	45

注：单元格中的数据是某个因变量的假设值。

然而，整体水平的稳定性和工作满意度的年龄模式也迫使我们考虑这样的事实，即工作的条件在很多方面都有所变化(可能变得更好)，而近些年连续进入职场的年轻人世代明显对于工作条件和回报有着更高的期望。因此，这个稳定的年龄模式似乎部分反映了相互抵消的时期和世代效应。年龄和工作满意度之间关系对线性的偏离，即每一年年龄的增

长带来的工作满意度的增加在中年阶段比年轻和年长阶段较小，则显示了年龄或组成效应，这使得三个或四个变量的解释对这个例子更加适合。

表1.2、表1.3和表1.4中因变量的变异都是线性的，但是年龄、时期和世代的非线性效应并没有像线性效应那样混淆。例如，表1.6中因变量指定为年龄变异的非线性模式。对这个数据仅有的解释为，在没有组成效应存在的情况下，它们反映了非线性的年龄效应，因为年龄变异非线性模式横跨了不同的世代和时期。如果时期的非线性模式横跨了不同的年龄和世代，或者世代的非线性模式横跨了不同的年龄和时期，则对数据的解释也非常清楚。如果这个数据反映了超过一种效应，那么对效应的估计当然无法通过对列联表的简单观测来完成。但是，正如我将在下一章解释的，如果所有的效应都是非线性的，那么在统计上可以依据一定的精度估计年龄、时期和世代效应。但是，因变量分年龄、时期和世代对线性变异的偏离并不意味着这些效应可以在统计上分离出来。如果变异的任何主要组成是线性的，那么对这些效应的统计分离是不可能的。

尽管识别问题涉及对数学定理的基本且广泛的理解以及世代分析之外的其他研究方法(例如，估计流动或者丈夫—妻子间不同的效应)，但直到20世纪60年代，这个问题才被大部分社会科学家意识到。当时小休伯特·M.布拉洛克(Hubert M. Blalock, Jr.)发表了一系列关于这个问题的文章(Blalock, 1966、1967)。直到1973年才出现了一篇之后被广泛引用的文章讨论了世代分析中的识别问题(Mason, Mason, Winsborough & Poole, 1973)。这篇杰作

的作者用简单的语言解释了涉及的逻辑和数学问题。但不幸的是，这篇文章对利用统计方法解决识别问题提出了不切实际的期望，因此，它也造成了试图分离世代数据中这些混合效应的一系列错误的尝试。其中一些将在下一章进行介绍。

第2章

估计年龄、时期和世代效应的策略

第1节 梅森、温斯伯勒和普尔等人的方法

在世代分析中用于估计年龄、时期和世代效应最常用的方法就是梅森等人在1973年的经典文章中所介绍的方法。这个方法和它的一些变形作为有力的方法被一直沿用下来。

梅森等人的方法的吸引力在于其简单易行。年龄、时期和世代都被记录为一系列虚拟变量,每一个虚拟变量代表5年或10年。当一系列变量作为预测变量被纳入回归,则其中的一个变量必须被忽略以使程序能够运行。当将年龄、时期和世代虚拟变量包含进来的时候,要使程序可以运行,我们就必须忽略一个额外的虚拟变量,它可以来自年龄、时期或者世代。忽略这个引入分析的附加变量与忽略一系列变量中的两个变量的效果是一样的,并且这一假设往往与真实状况并不相悖,尤其当被忽略的两个虚拟变量代表的类别是相邻类别的时候。

上述方法当然不是唯一在分析中同时加入年龄、时期和世代作为预测变量而能够使程序运行的方法。有许多方法可以打破这些变量间的线性相关并使程序运行,例如对变量进行重新编码或者转换,即加以识别的限定。例如,年龄和世代可以作为连续变量而时期作为一系列虚拟变量加入回

归。然而当我们这么处理的时候，线性相关的问题也仅仅在统计模型中被打破，但在现实世界中并没有被打破，因此，估计出的效应并不具有实际意义。

因此当我们决定使用梅森等人的方法时，这种选择并非随意的。许多研究者相信，由于相邻的年龄、时期或者世代的类别的效应是相同的这一假设并不违背实际，因此选择任意两个相邻的类别就可以。但这是梅森等人从未提及的。

许多使用梅森等人的方法或其变形的研究者也忽略了文章中介绍的另一个重要部分，而这一部分内容由于出现在文章的脚注中，所以很难被注意到：

> 这些纯粹的效应（用于解释这一方法的假设效应）的形式是刻意设定为非线性的……我们将数据构造成这种形式的原因在于，完全的线性效应是难以被解释清楚的，而且在世代分析中，根据纯粹效应的变量和两个剩余变量估计的结果是一样的。(Mason et al. , 1973: 248)[1]

换句话说，这种方法仅适用于估计非线性效应。

当效应是线性的时候，这种方法的弱点就在表 2.1 中显现了，其中年龄、时期和世代可根据基于不同简化假设的模型来估计出其效应。为了使分析更简单易懂，我构造了一个数据集，它符合每 10 年对因变量有 5 点的正向时期效应以及 5 点的负向世代效应这一假设，且在 1950 年时，20—29 岁之间的受访者的因变量值为 50。表 1.2 显示了构造后的数据结果。对于这个模拟的实验，我知道计算这些效应可以利用梅森等人的方法并观察其运行状况，因此我使用了 4 个具有不

同简化假设(识别限定)的模型。表 2.1 是模型的估计结果。

表 2.1 利用梅森等人的年龄—时期—世代模型估计反映在表 1.2 中的效应(未标准化的回归系数)

变量	模型	1	2	3	4
常数项		50.0	50.0	25.0	28.7
	年龄				
	20—29	a	a	0.0	a
	30—39	5.0	5.0	0.0	a
	40—49	10.0	10.0	a	3.6
	50—59	15.0	15.0	a	6.1
	60—69	20.0	20.0	0.0	8.6
	70—79	25.0	25.0	0.0	11.4
	时期				
	1950	a	0.0	a	a
	1960	0.0	0.0	5.0	a
	1970	0.0	0.0	10.0	3.6
	1980	0.0	a	15.0	6.1
	1990	0.0	a	20.0	8.6
	2000	0.0	0.0	25.0	11.4
	世代(出生年份)				
	1871—1880	0.0	0.0	50.0	25.0
	1881—1890	0.0	0.0	45.0	23.9
	1891—1900	0.0	0.0	40.0	21.5
	1901—1910	0.0	0.0	35.0	19.1
	1911—1920	0.0	0.0	30.0	16.8
	1921—1930	0.0	0.0	25.0	14.0
	1931—1940	0.0	0.0	20.0	11.8
	1941—1950	0.0	0.0	15.0	9.1
	1951—1960	a	0.0	10.0	6.5
	1961—1970	a	0.0	5.0	3.9
	1971—1980	0.0	a	a	a

注:a. 参照类,其值为 0。

除非相邻两个年龄层的效应是一样的(这样可以获得正确的估计),否则这个方法给出的结果大体上是不正确的。如果两年或者两个世代效应相同的假设是适用的,那么这个

方法产生了一个单变量(完全的年龄效应)解决方法,同时,假设这一效应在两个年龄层和两年间是相同的,则会产生一个三变量的解释。数据本身并不能提供应该使用哪个相同假设的线索,因为每一个假设对于现实的扭曲程度是相同的(当然,相比一个假设,两个假设引入了更多的扭曲)。显然,能够利用模型获得正确的估计取决于对哪一种解释的预先了解是正确的。这种方法并不能告诉研究者究竟是变量中哪一种组合的解决方式最能精确地代表现实。第 1 章中介绍的其他方法同样不能解决这个问题。但是,由于非线性的年龄、时期和世代效应彼此之间的混合与线性的不同,人们可能会认为,这些效应可以由梅森等人的方法来很好地估计。另一个模拟的实验可以提供相关的证据。这一次,我假设非线性的纯粹年龄效应,其中,因变量在 20—29 岁群体中的值为 50,且此后每 10 年的衰老具有的效应分别为＋2、＋10、0、－12 和－5,直至 70—79 岁年龄段。这种模式的效应将产生表 1.6 中的数据。我再一次使用梅森等人的方法,用 4 种不同模型来估计反映在数据(表 2.2)中的效应。对于模型 1,其背后的假设为 40—49 岁的效应和 50—59 岁[①]的效应是一样的,而这是符合现实的,所以这个模型获得了准确的估计。对于模型 2,其背后的假设为,任意两个世代的效应是相同的,而这也是正确的,因为没有世代效应,从而使得 APC 的所有估计都是正确的。对于模型 3,其背后的假设为,任意两个时期的效应是相同的,同样,正确的简化假设使得对年龄、时期和世代效应的估计都是正确的。对于模型 4,

① 原书此处为 30—39 和 40—49,应该是笔误。——译者注

其背后的简化假设为,20—29岁和30—39岁的效应是一样的,但这并不正确,尽管误差并不大,但是可能导致与真实效应完全相反的估计。鉴于一个变量(纯粹年龄效应)的方法是正确的,这个模型显示了一个三变量模式的效应。

表 2.2 利用梅森等人的年龄—时期—世代模型估计反映在表 1.6 中的效应(未标准化的回归系数)

变 量	模 型	1	2	3	4
常数项		62.0	50.0	50.0	60.0
	年 龄				
	20—29	−12.0	a	a	a
	30—39	−10.0	2.0	2.0	a
	40—49	a	12.0	12.0	8.0
	50—59	a	12.0	12.0	6.0
	60—69	−12.0	0.0	0.0	−8.0
	70—79	−17.0	−5.0	−5.0	−15.0
	时 期				
	1950	a	a	a	a
	1960	0.0	0.0	a	2.0
	1970	0.0	0.0	0.0	4.0
	1980	0.0	0.0	0.0	6.0
	1990	0.0	0.0	0.0	8.0
	2000	0.0	0.0	0.0	10.0
	世 代 (出生年份)				
	1871—1880	0.0	0.0	a	20.0
	1881—1890	0.0	0.0	0.0	18.0
	1891—1900	0.0	0.0	0.0	16.0
	1901—1910	0.0	0.0	0.0	14.0
	1911—1920	0.0	0.0	0.0	12.0
	1921—1930	0.0	0.0	0.0	10.0
	1931—1940	0.0	0.0	0.0	8.0
	1941—1950	0.0	0.0	0.0	6.0
	1951—1960	0.0	0.0	0.0	4.0
	1961—1970	0.0	a	0.0	2.0
	1971—1980	a	a	0.0	a

注:a. 参照类,其值为0。

从中获得的经验是显而易见的:哪怕效应的非线性模式与表1.6中所示的一样简单,但只有当简化假设是正确的时候,梅森等人的方法才能够获得正确的估计。然而,一个研究者很少能够确信某一个简化假设是一点都不含有误差的。

当效应的非线性模式更复杂的时候,即包括两个或全部APC变量的非线性模式时,梅森等人的方法或者其他的统计模型都不能充分解决问题。哪怕非线性效应彼此之间并不像线性模式中那样相互混淆,但它们之间仍然可能存在APC变量之间的交互。例如,可能非线性的年龄效应在统计分析中反映为时期和世代之间的交互,但并非所有的时期—世代交互都是年龄效应。表1.6中的非线性变异的简单模式使我们很容易判断出变异反映了年龄效应,但是更复杂的数据则无法通过简单的观察得出结论。

没有考虑到APC之间的交互是使得梅森等人的方法并不像很多研究者相信的那样有用的另一个原因。所有关于这个方法的变形都基于年龄、时期和世代效应是叠加的这一假设,也就是说,年龄效应对于所有的时期和世代都是相同的,时期效应对于所有的年龄和世代都是相同的,以及世代效应对所有的年龄和时期都是相同的。这个假设可以进行检验,但往往并不符合现实。当态度和行为的变化发生在成年人群体中时,变化往往在较年轻的人中更为突出。数据本身并不能分辨年龄—时期和世代—时期的交互,但是有理论表明,随着年龄的增大,成年人由于刺激而产生的变化逐渐变小(Glenn, 1974、1980; Alwin, Cohen & Newcomb, 1991)。态度的变化往往随着年龄的变化而不同,无论是更加激进还是更加保守,这也是我下一章即将讨论的问题。此

外，一些年龄效应(尤其是那些最近发生的)与社会角色而非生物衰老相关的变化，也造成世代间的不同，例如对中老年人性生活活跃度的期望有所上升。世代效应也可能随着年龄而变化，例如一小部分世代群体在经济和职业上的优势随着他们年龄的增长而扩大，却又随着更年轻世代群体的挑战而慢慢减小。

通过对表 1.1 中女性结婚率的数据进行检验，可以揭示出叠加的 APC 模型的不完善之处以及利用附加信息解释世代数据的需求。这个数据中存在大量的交互。尽管结婚率在所有年代都随着年龄负向变化，但用年龄对 1968 年和 1998 年结婚率进行的回归结果显示，其系数从－1.012变为－2.93。此外，虽然在 1968 年结婚率最高的是在 25—34 岁年龄段，但在其后的时期，结婚率最高的年龄段变为35—44 岁。同时，用年份对结婚率进行回归，其回归系数从25—34 岁年龄段的－0.696 变为 65—74 岁年龄段的0.299。1968 年处于 25—34 岁年龄段的世代在迈向 35—44 岁年龄段时，经历了结婚率的下降，但是经过同样的衰老期，1978 年处于 25—34 岁年龄段的世代并没有经历结婚率的下降，而 1988 年处于 25—34 岁年龄段的世代则经历了结婚率的上升。

对这些强调交互的效应进行统计建模是困难的，哪怕识别问题并不会产生干扰。然而，对于熟悉美国近些年来婚姻、离婚和寿命的家庭人口学者来说，这种数据并非十分神秘。相关的趋势为：(1)从 20 世纪 70 年代末至 20 世纪 80 年代，典型的初婚年龄有显著的上升；(2)从 20 世纪 60 年代中期至 20 世纪 80 年代，离婚率有显著的上升，而持久的婚姻

仅占很小一部分;(3)中老年死亡率从 20 世纪 80 年代之后又持续下降;(4)进入较老年龄段的世代具有最高的终身结婚率。如果不了解这些趋势,对表 1.1 中的数据进行统计处理则并不能带来对数据复杂模式的深入理解。

第2节 | 中村的贝叶斯方法

梅森等人的世代分析方法只有在简化假设(识别限定)是可以随意选择的情况下才可以随意地作为一种常规方法来运用,但这是一种对方法的滥用。然而,其他世代分析方法则希望建立一套程序化的应用,也就是说,无论理论或者附加信息预测出怎样的年龄、时期和世代效应,这种方法都可以以相同的方式进行应用。

这些方法中最著名的可能是来自日本统计学家中村的研究(T. Nakamura, 1982、1986),并由佐佐木和铃木介绍给美国社会科学家(Sasaki & Suzuki, 1987)。这个方法不变的简化假设就是连续的参数是逐步变化的。更具体地说,这个假设就是在年龄、时期和世代中,相邻的比分散的更为接近。佐佐木和铃木给出了进一步的解释:

> 对于一个社会现象,较短间隔内进行调查产生的百分比变化比较长间隔内进行调查产生的百分比变化要小。例如对宗教的热忱度,20岁年龄段和30岁年龄段群体之间的差距大致上比20岁年龄段和40岁年龄段之间的差距小。(Sasaki & Suzuki, 1987:764)

佐佐木和铃木承认，连续的参数是逐步变化的这一假设并非永远正确。研究者只需要观察 APC 变量中任意两者与因变量之间的关系是否是单调即可。如果确实如此，或者只有很少的违背之处，则这个假设基本上是正确的。

如果佐佐木和铃木对他们观点的阐述是有效的，那么这个方法的确是有用的，因为逐步变化这个假设可以实证地进行检验。佐佐木和铃木将他们的方法应用到与表 1.4 非常类似的世代数据中(Sasaki & Suzuki, 1989)，所有由年龄和时期导致的变异不仅是单调的，而且是完全线性的。尽管有无数种可能的解释且其中一些对于特定的因变量是合理的，但是他们总结道，数据只反映了世代效应。这显示出这个方法往往选择所有可能的解释中最简单的，至少是像这个方法所限定的简约性一样。这种解释可能通常是正确的，但仅仅是"通常"，这对于建议研究者使用这个方法且接受这个估计是正确的理由并不充分。同样，相比于考虑了更多信息而并不十分有力的方法得出的结论，中村的估计方法并不更加准确。

显然，对中村的方法应该持怀疑的态度，但是如果估计是根据理论和附加信息或者是对数据模式的评估信息而获得的，那么便可能是有效的。例如表 2.3 中，佐佐木和铃木对荷兰关于无宗教归属人士的数据进行的分析(Sasaki & Suzuki, 1987)。在一个独立的表中，作者指出，在所有的荷兰成年人口中，无宗教归属人士从 1899 年的 1.8%单调地增长至 1969 年的 23.0%。然而从表 2.3 的数据中可以看出，时期效应对于无宗教归属有很强的影响。当这样的时期影响持续时，便会出现一个与全部人口中变化方向一样的世代

间的变化趋势。因此,任何检视表 2.3 的分析者都能发现,该表显示出了数据中较大的正向时期和世代效应。尽管非宗教归属在表中每个世代内都经历了超过一个时期的上升,但并没有其他来源的理论或证据显示,随着年龄的增长宗教归属感会下降。因此,世代间的变化很大程度上应该是时期效应。

表 2.3 荷兰无宗教信仰人口百分比(根据年龄和年份)

年份	年龄					
	20—29	30—39	40—49	50—59	60—69	70—79
1899	2.4	2.2	1.7	1.2	0.9	0.6
1909	5.8	5.6	4.2	3.2	2.4	1.6
1919	8.4	8.9	7.1	5.3	3.9	2.7
1929	15.0	16.2	14.1	11.0	8.1	5.8
1939	16.9	18.3	17.0	14.6	11.2	8.0
1949	18.0	19.5	19.0	17.5	14.3	10.4
1959	18.3	20.2	19.8	19.2	17.8	14.2
1969	24.7	23.0	23.3	23.1	21.4	19.2

资料来源:佐佐木和铃木书中的表 4(Sasaki & Suzuki, 1987)。

表 2.4 中村的贝叶斯方法和一般最小二乘法估计得出的荷兰非宗教归属的年龄、时期和世代效应

	中村的贝叶斯估计	一般最小二乘回归
年龄		
20—29	−0.715	a
30—39	0.0641	a
40—49	0.0615	a
50—59	0.0473	a
60—69	−0.0063	a
70—79	−0.0950	a
时期(年份)		
1899	−1.2301	a

续表

	中村的贝叶斯估计	一般最小二乘回归
1909	−0.5361	1.352
1919	−0.2775	2.269
1929	0.2625	6.385
1939	0.3305	7.406
1949	0.3985	8.042
1959	0.4355	8.732
1969	0.6177	11.835
世代(出生年份)		
1820—1829	−1.2967	−5.686
1830—1839	−1.0314	−7.712
1840—1849	−0.7699	−5.393
1850—1859	−0.4908	−5.137
1860—1869	−0.2230	−4.208
1870—1879	0.0447	−2.578
1880—1889	0.2913	a
1890—1899	0.4545	2.286
1900—1909	0.5195	3.814
1910—1919	0.5575	4.536
1920—1929	0.5754	4.678
1930—1939	0.5539	4.081
1940—1949	0.7742	6.579

资料来源:格伦书中的表3(Glenn, 1989)。

表2.4显示了中村对数据中效应的估计,且估计结果精确地反映了理论、附加信息和对数据的非正式评估所指示的效应模式。在这个例子中,这个方法显然非常实用。另一方面,研究者也可以通过简单的假设来估计不存在年龄效应的情况而获得类似的结果,将时期和世代都转换为虚拟变量,并使用最小二乘回归来估计时期和世代对非宗教归属的效应(参见表2.4中最后一列)。从这个例子中我们可以看出,

复杂的中村方法并不比简单的回归分析有效,尽管在其他的例子中也许是这样。

中村的贝叶斯方法并没有被广泛地使用(至少在美国如此)。佐佐木和铃木认为,这种方法为世代分析中的识别问题提供了一个全面的解决方法,这一观点已经大致上被否定了。然而,这个方法有时可以对效应提供可信的估计,它也尽可能努力地成为分析世代的全面方法。因此,尽管中村方法并没有成功,但仍有许多后续的研究希望提供一个对世代分析普遍适用的统计方法。

第3节 后续的探索

在本书即将完成的时候,有很多正准备出版的著作提出并阐释了一种能够将 APC 效应模型化的新的统计方法。由于这些著作仍处在“请勿引用”的阶段,我并不能总结它们的内容或者介绍其作者。但至少有一些作者将其新的方法称做“解决 APC 难题的全能方法”以招徕读者,他们宣称,当数据为短时期多次收集获得的、可比较的重复截面数据时,应用这个方法将获得有效的估计。当这些方法利用已经发表的文章进行解释时,我们并不能要求其改变分析,并考虑到理论和附加信息以及在 APC 变量线性模式中因变量方差的含混性。换句话说,至少有一些方法的鼓吹者号称他们可以解决逻辑上不可能的问题。

当然,如果“通常”能够获得接近正确的估计,研究者能够使用理论和附加信息来评价估计的可信度,且他们对所获得的结论持探索性态度,那么这种方法也许可以被证明是有效的。

我对那些想要使用这种方法的研究者的建议是,不妨等到它已经被不同情境下模拟的实验状况完全地检验后再使用。或者,研究者可以自己进行这类模拟。当模拟完毕后,研究者应该牢记,几个对效应成功估计的例子并不能够证明方法是普遍适用的。此外,在模拟实验中,也应该将复杂的效应混合情况考虑进去。

第4节 | 年龄—时期—世代—特征模型

对世代效应研究的最新进展是年龄—时期—世代—特征(APCC)模型(O'Brein, 1989、2000; O'Brein, Stockard & Isaacson, 1999)。这些模型包括一个或一个以上的"世代特征",例如世代规模或者对随世代群体而变化的家庭结构的测量。尽管不能控制年龄和时期而单独让世代变化,但是当世代的特征与世代的相关程度不太高的时候,是可以控制年龄和时期而让世代的特征变化的。如果研究的目标仅仅是为了估计选择出来的世代特征的效应,而非年龄或时期的效应,并且世代特征与世代之间并没有强烈的线性关系,那么这些模型就可以被有效地应用。但是需要记住的是,这些模型并不是真正的 APC 模型(尽管有时它们也用来表示 APC),且它们并不能够解决年龄—时期—世代这个难题。

某个因变量的世代效应不可能仅仅是因为模型中包含的世代特征所导致的,而模型估计中的世代效应也与年龄和时期效应混淆不清,这样做可能导致模型估计出的时期与年龄效应没有意义。此外,当世代特征与世代的线性相关较严重时,世代特征的效应也容易与时期和年龄效应相混淆。幸运的是,在最近的美国数据中,世代规模与世代群体的出生日期线性相关较低,因此利用 APCC 模型估计世代规模的效

应也会较为准确。然而一些其他的世代特征，例如婚外所生的婴儿比例和在未成年期经历了父母离婚的世代成员的比例，则与世代有单调的线性关系（尽管并非完全线性）。这些特征都被应用在APCC模型中，但是它们与世代之间严重的线性关系会导致对其效应的估计不那么准确。[2]

当然，研究者可以同时构想年龄—特征—时期—世代模型和年龄—时期—特征—世代模型。尽管我没有看到此类方法的专有术语，但是在某些研究中，时期、世代和一些特征与年龄是具有一定程度相关的，且它们同时被当做自变量，此外，还有些研究中使用了可以被称做"年龄—时期—特征—世代模型"的方法。与APC模型一样，这些模型同样具有优点与缺点，因而需要尽量避免自变量与忽略的APC变量具有的很强的线性关系，而这样也导致对包含APC变量的估计没有意义。

第5节 | 对APC效应估算的非正式方法

不能通过统计上的APC模型对年龄、时期和世代效应作出准确估计这一事实并不意味着我们要放弃对这些效应的区分。一旦不能精准和绝对确信地分离这些效应,那么利用理论、附加信息、常识以及多种统计分析来进行判断就是合理的。但是并没有公式或者手册适用于区分所有情况下的这些效应。在少数情况下,只需对交互数据进行目测即可,这时候我们会发现,因变量的变化与表1.6中的情形非常相似。同样,对于自变量在世代间变化的非线性模式在不同年龄和时期是一致的情况下,或者不同时期间变化的非线性模式在不同年龄和世代是一致的情况下,对此都只有一种可能的解释。但是对这些效应作出合理的判断往往是困难的,且需要研究者具有巧妙的心思、深谋远虑和良好的判断。

为了使年龄、时期和世代效应的估计获得足够确信的结论,我们往往需要以很多种方法来审视数据,并引入很多种附加信息(参见Abramson & Englehar, 1995; Alwin, 1991; Converse, 1976; Glenn, 1994、1998)。而从简单到复杂的各种统计方法,也可能起到一定的作用。

一种简单而有效的方法就是对介入年龄、时期或者世代

和因变量之间的变量进行统计控制。我们以美国为例来解释这个问题，在农村长大的人口比例随着不同的世代群体的变化而有很大的改变，那么，通过控制出生人口所在社区的规模来移除任何由年龄导致的自变量变异，则可能是世代效应而非年龄效应。例如，同胞数量或者父母是否离异这样的未成年时期的背景变量则可能同时介入世代与一些因变量之间。教育程度是另一个与世代相关的变量，但是对控制它所获得结果的解释必须十分小心，因为教育对于不同的世代自然会有不同的效应（Alwin，1991；Glenn，1994）。例如，婚姻状况、父母状况、就业状况以及生理健康都可能介入年龄与各种因变量之间；富裕程度、失业率和政治大环境都是与时期相关的变量，而在世代分析中对其进行控制也较为有效。

由于非正式的世代分析方法必须根据所面对的特殊问题进行应用，我并不能对所有可以被应用的有效技术一一进行介绍和解释。但是，这里我举一个以个人幸福水平为因变量的例子来解释一些有用的策略。

第6节 实例:一项有关个人幸福感的世代研究

这个研究的目的在于尝试性地回答这些问题:第一,在美国社会中,成年人年龄的增长对个人幸福感有何影响?第二,这些影响在男性和女性间有什么差别?

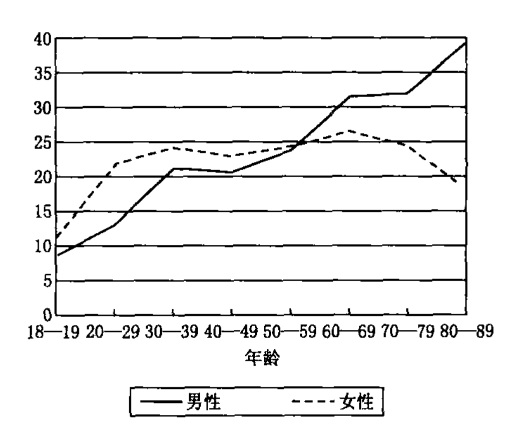
这个研究的数据来源是芝加哥大学的国家民意研究中心1972年至2002年间进行的美国综合社会调查(GSS),这个调查从1972年至1978年每年进行一次,1980年进行一次,1982年至1991年每年进行一次,1993年、1994年、1996年、1998年、2000年和2002年也各进行一次。在每次的调查中,受访者都会被问到如下问题:

> 综合来看,你最近的感受是非常幸福、一般幸福还是不怎么幸福?

在研究中,这个定序的三点度量往往被作为定距变量进行处理,这里我进行的回归分析也同样如此。然而,我使用了另外两个测量对汇总层次进行分析:(1)幸福指数,即用选择"非常幸福"的受访者比例减去选择"不怎么幸福"的受访者的比例;(2)对于选择"非常幸福"与选择其他的受访者进

行一个简单的二分。

把从1972年至2002年的数据合并在一起所得的图2.1显示了幸福指数(HI)与年龄的关系(男女分别计算)。



注:a. 选择“非常幸福”的受访者比例减去选择“不怎么幸福”的受访者的比例。

资料来源:美国综合社会调查1972年至2002年的合并数据(Davis et al., 2002)。

图2.1 根据年龄和性别绘制的幸福指数[a]

对于男性,年龄与幸福指数的关系是单调递增的;对于女性,幸福指数值在中年的时候是最高的,在更高年龄时降低,而在青年时期是最低的。在50岁以前,相对于男性,女性有较高的幸福指数;但是50岁之后,这个关系则正好反了过来。

正如读者所知,幸福指数与年龄的关系也许被年龄、世代和删减(组成的)效应所混淆。对这个数据的第一反应的解释可能是男性和女性的幸福感是如何随着年龄而变化的,但是也有另外的解释。例如,男性幸福指数与年龄的单调正相关也可能由于每一个成年的世代群体都比之前的世代群体更加不幸福而导致的。此外,男性与女性的差别与年龄之

间的关系也许是由于这些年轻女性在美国社会所获得的进步导致的,这也造成了与较老的女性出生群体相比,较年轻的女性出生群体与男性之间的差距较小。

表 2.5 根据年龄、时期和性别的幸福指数[a](美国,括号中为人数)[b]

年龄	1972—1982		1983—1992		1993—2002	
	男性	女性	男性	女性	男性	女性
18—19	4.4 (250)	9.0 (211)	11.9 (168)	12.6 (151)	11.2 (143)	13.4 (134)
20—29	8.8 (1059)	21.7 (1195)	17.7 (1447)	23.4 (1679)	13.6 (1261)	19.5 (1364)
30—39	19.3 (1218)	26.9 (1495)	21.9 (1343)	22.0 (1680)	22.2 (1416)	23.5 (1666)
40—49	19.6 (1158)	26.2 (1366)	22.1 (1074)	23.6 (1323)	19.8 (1348)	19.5 (1617)
50—59	23.4 (1183)	25.8 (1319)	23.3 (803)	20.1 (935)	25.2 (917)	25.8 (1078)
60—69	28.6 (829)	25.6 (906)	32.7 (707)	27.9 (922)	34.4 (591)	25.8 (704)
70—79	30.5 (471)	25.6 (510)	30.7 (407)	26.1 (583)	34.7 (427)	21.6 (555)
80—89	33.1 (118)	18.3 (147)	29.7 (128)	24.2 (223)	30.1 (146)	11.9 (252)
总计	19.1 (6958)	24.5 (7869)	22.7 (6069)	23.3 (7496)	22.3 (6249)	21.7 (7370)

注:a. 选择"非常幸福"的受访者比例减去选择"不怎么幸福"的受访者的比例。

b. 受访者的人数是根据权重近似获得的。数据的权重是家庭中成年人的比例除以 GSS 家庭中的成年人平均数量。

资料来源:美国综合社会调查 1972 年至 2002 年的合并数据(Davis et al., 2002)。

当合并的截面数据包括了几年甚至几十年时,我们必须观察获得的关系是否在这段时间内是稳定的。如果不是,那么其中的变化则可以显示出数据应该进行怎样的解释。对数据的目测观察并不能发现年龄与幸福指数之间的变化模

式(无论男性还是女性),这一事实也同样被分性别和时期、根据年龄对于(幸福感)3 点度量的回归结果所验证。男性在 3 个时期的非标准化回归系数分别为 0.004、0.003、0.004,且每个系数的双尾检验在 0.01 水平上均显著。而类似的女性系数则为 0.001、0.001、0.000,但没有一个在统计上是显著的。因此,对合并数据进行解释时,不考虑年龄与幸福指数之间的关系在 1972 年至 2002 年间的变化是合理的。

表 2.6　美国人分时期、性别选择"非常幸福"[a] 人数的百分比(括号中的为人数)[b]

	男性	女性	男女差异
时期			
所有年龄			
1972—1982	33.2 (6973)	36.6 (7905)	−4.4***
1983—1992	32.8 (6084)	33.6 (7523)	−0.8
1993—2002	32.4 (6256)	32.7 (7388)	−0.3
变化			4.1***
年龄(18—34 岁)			
1972—1982	24.7 (2641)	34.8 (2935)	−10.1***
1983—1992	28.3 (2286)	31.6 (2712)	−3.3*
1993—2002	27.6 (2113)	31.4 (2328)	−3.8**
变化			6.3***

注:a. 其他的受访者选择的是"一般幸福"和"不太幸福"。
b. 受访者的人数是根据权重近似获得的。数据的权重是家庭中成年人的比例除以 GSS 家庭中的成年人平均数量。
* 表示 $p<0.05$ (双尾检验); ** 表示 $p<0.01$ (双尾检验); *** 表示 $p<0.001$ (双尾检验)。
资料来源:美国综合社会调查 1972 年至 2002 年的合并数据(Davis et al.,2002)。

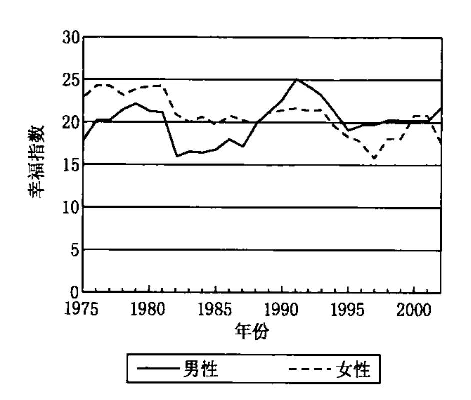
然而,1972 年至 2002 年间的变化是很重要的,且必须考虑进来。例如,如果利用世代效应来解释男性的幸福指数与年龄之间的正相关是正确的,那么男性整体的幸福程度应该在 1972 年至 2002 年间呈下降趋势,且在最年轻的成年人处下降最快。尽管不能完全拒绝此解释,但是此解释遭到表 2.6和表 2.7 所显示数据很强的质疑。这些数据显示了男性的平均幸福程度在上升,尤其在 18—34 岁的年龄阶段。同样,如果女性在较低年龄具有较高幸福水平和男性在较高年龄具有较高幸福水平完全或大部分是由于世代效应所导致的,那么女性的幸福水平应该在 1972 年至 2002 年间不断逼近男性,尤其在年轻成年人中。然而,图 2.2 和表 2.6、表 2.7显示出了相反的趋势。男性不断逼近女性,且与整体相比,这种变化在年轻成年人间更加明显。这种解释是基于没有严重的时期效应抵消了世代效应这一假设之上。这是一个合理的假设,因为时期与世代对整体趋势的影响效应往往(并非始终)相互加强,而非彼此抵消。

表 2.7 对美国人幸福度[a] 根据年(1972—2002 年)和性别回归的结果

	男性	女性	男女差异
所有年龄	0.002***	−0.001	0.003***
年龄(18—34 岁)	0.003***	−0.001	0.004**

注:a. 幸福水平是根据 3 点进行度量的,且处理为定距变量。
** 表示 $p < 0.01$ (双尾); *** 表示 $p < 0.001$ (双尾)。
资料来源:美国综合社会调查 1972 年至 2002 年的合并数据(Davis et al., 2002)。

图 2.1 显示出的幸福指数与年龄之间的关系反映了年龄效应,这也可以由反映幸福水平随着世代群体年龄在 1972 年至 2002 年间增长趋势的数据加以补充说明。追踪世代群



注：a. 选择“非常幸福”的受访者比例减去选择“不怎么幸福”的受访者的比例。

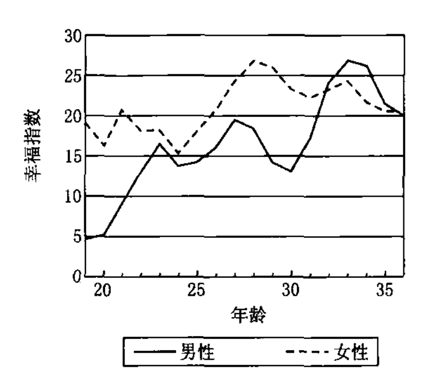
b. 所指年份为 4 年一轮的最后一年。

资料来源：美国综合社会调查 1972 年至 2002 年的合并数据（Davis et al.，2002）。

图 2.2　根据时期和性别的幸福指数[a]（4 年一轮的均值[b]，美国）

体间的趋势往往通过选择一个较宽的出生年份区间的出生群体（例如一个 5 年或 10 年的出生群体开始），并从数据中的第一年到最后一年对他们进行追踪。这个步骤的一个缺点在于，对世代中的年轻成员与较老成员进行的追踪并非经过同样的年龄。例如，我们选择 1972 年年龄为 20—29 岁的世代，并追踪他们至 2002 年，最老的成员是从 29 岁追踪至 59 岁，而最年轻的成员则是从 20 岁追踪至 50 岁。对某些研究而言，这个步骤已经足够了，但是当我们想要更清楚地了解在某个特殊的年龄发生了什么时，则需要对每个 5 年或 10 年的世代在同样的年龄进行追踪。当然，在进行此步骤时，不同世代的成员是在不同的时期进行追踪的。如果数据被限制为能够包含所有世代成员的时间点（这是为了使得样本规模足够大以获得可靠的估计），则时间的跨度可能变小。在 1972—2002 年的

GSS数据中,第二种方法只能追踪20年的世代,而非30年。



注:a. 选择"非常幸福"的受访者比例减去选择"不怎么幸福"的受访者的比例。

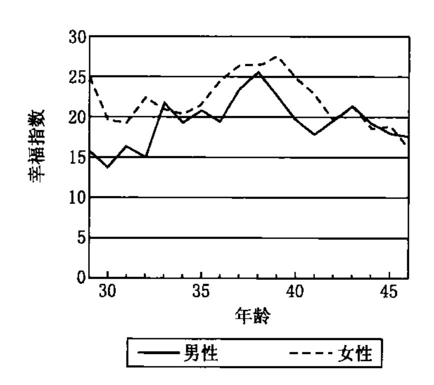
b. 所指年龄为3年一轮中的中间年龄。

资料来源:美国综合社会调查1972年至2002年的合并数据(Davis et al. , 2002)。

图2.3　1955—1964年世代的幸福指数[a]
(年龄从18岁至37岁,3年一轮的平均值)[b]

在图2.3至图2.7中,我们显示了幸福指数在世代内部的趋势,共5个10年世代群体,从1915—1924年世代到1955—1964年世代,综合起来从18岁追踪至77岁。如果对图2.1中数据的年龄效应解释是正确的,那么在每个世代中,男性应该经历幸福指数的上升,且男性的趋势应该比女性的更加令人满意。图2.3至图2.7和表2.8也大致符合这一预测。事实上,预测出的变化比图2.1中显示将导致产生年龄与幸福指数之间关系的变化更大。截面数据中男性幸福指数(3点度量)根据年龄的回归系数(非标准化)为0.002(图2.7),根据世代内趋势数据获得相应的每年的变化为0.006(世代的均值根据表2.8中的数据计算获得)。此外,截面数据中男性与女性回归系数之间的差别为0.003

（表2.7），但是对男女之间幸福指数差异预测的世代内每年平均的变化为0.008（根据表2.8中的数据计算获得）。

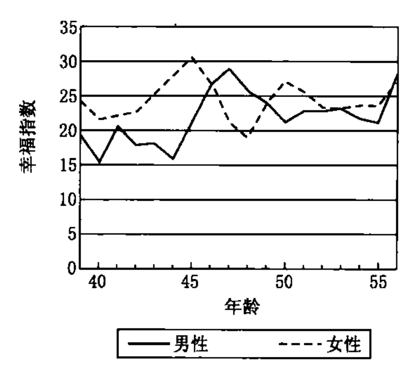


注：a. 选择"非常幸福"的受访者比例减去选择"不怎么幸福"的受访者的比例。

b. 所指年龄为3年一轮中的中间年龄。

资料来源：美国综合社会调查1972年至2002年的合并数据(Davis et al.，2002)。

图2.4 1945—1954年世代的幸福指数[a]
(年龄从28岁至47岁，3年一轮的平均值)[b]

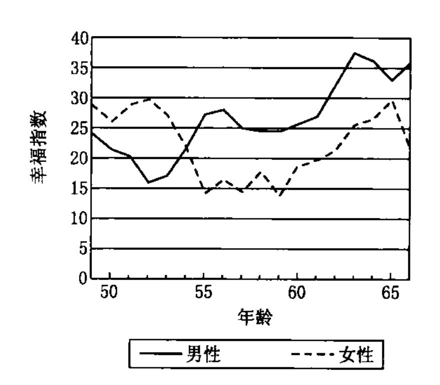


注：a. 选择"非常幸福"的受访者比例减去选择"不怎么幸福"的受访者的比例。

b. 所指年龄为3年一轮中的中间年龄。

资料来源：美国综合社会调查1972年至2002年的合并数据(Davis et al.，2002)。

图2.5 1935—1944年世代的幸福指数[a]
(年龄从38岁至57岁，3年一轮的平均值)[b]

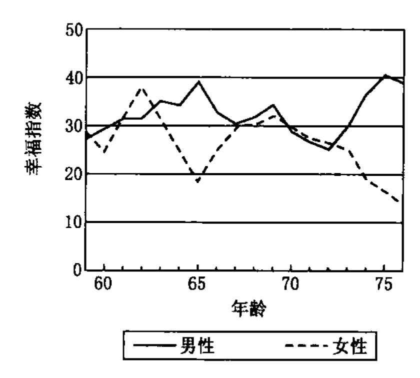


注:a. 选择"非常幸福"的受访者比例减去选择"不怎么幸福"的受访者的比例。

b. 所指年龄为3年一轮中的中间年龄。

资料来源:美国综合社会调查1972年至2002年的合并数据(Davis et al.,2002)。

图2.6 1925—1934年世代的幸福指数[a]
(年龄从48岁至67岁,3年一轮的平均值)[b]



注:a. 选择"非常幸福"的受访者比例减去选择"不怎么幸福"的受访者的比例。

b. 所指年龄为3年一轮中的中间年龄。

资料来源:美国综合社会调查1972年至2002年的合并数据(Davis et al.,2002)。

图2.7 1915—1924年世代的幸福指数[a]
(年龄从58岁至77岁,3年一轮的平均值)[b]

表2.8 20年后对美国5个10年一组的出生群体的幸福感根据年龄分性别进行回归(非标准化)

世代（年龄段）	男性	女性	男女差异
1915—1924（58—77）	0.004	−0.006*	0.010*
1925—1934（48—67）	0.009**	−0.004	0.013**
1935—1944（38—57）	0.005	0.000	0.005
1945—1954（28—47）	0.002	−0.003	0.005
1955—1964（18—37）	0.009***	0.003	0.006*

注：幸福水平是根据3点进行度量的，且处理为定距变量。
*表示 $p < 0.05$（双尾）；**表示 $p < 0.01$（双尾）；***表示 $p < 0.001$（双尾）。
资料来源：美国综合社会调查1972年至2002年的合并数据（Davis et al.，2002）。

对这个数据最合理的解释为，世代内的趋势是由年龄和时期效应混合导致的，时期效应明显是由时期的影响带来的，这在表2.6与表2.7中的年轻人的趋势中可以显示出来。换句话说，最近的时期效应使得男性的幸福指数相对女性升高，向逐渐老化的世代中的变化增添了同向的年龄效应。

表2.6与表2.7中显示的年轻人中的趋势表明了时期影响，但是它们同时也反映了世代效应。这些趋势在何种程度上可以被概念化为时期效应或者世代效应并不清楚，但是承认至少有一部分是世代效应，这对于解释年龄和幸福感的截面数据是十分必要的。例如，任何世代内幸福程度上升的趋势将会减少由于年龄正向效应导致的截面的年龄差异。

这意味着年龄对男性幸福感的正向效应往往比图 2.1 中截面数据所示的大。

美国在最近几十年来,男性的幸福感明显地随着年龄的增长而单调上升,而且由于他们变老,正向的年龄效应也比数据显示的大,正如图 2.1 中所表现的那样。女性的幸福感符合同样的单调趋势,因为随着她们年龄的增长,与男性几乎所有的年龄水平相比,女性的年龄效应较低,且在年纪较长时呈负向。

当世代分析的范围中包含较大年龄时,则应该考虑由于不同的死亡率而导致对因变量产生的组成效应。例如在这个研究中,不幸福的人可能比幸福的人死得更早,而这样就造成了数据上男性从 70 岁到 80 岁之间幸福感的大幅上升。同样,这种组成效应也可能发生在男性较年轻的阶段,并且也可能使得女性在较老时幸福感下降得没有那么严重。

并没有明确的证据能对这种状况发生的可能性进行评估。老年人所报告的幸福感与他们的长寿程度相关(Palmore & Jeffers, 1971),但这可能仅仅是由于人们在去世之前往往经历了健康程度的下降,而较差的自我感知健康程度与报告的幸福程度相关。如果是这样的话,那么死亡率对汇总层面的幸福感的正向效应则可能很大程度或者完全地被去世前健康程度下降的负向效应所抵消。在这里,我们需要但是难以获得的信息是幸福感与寿命在一生中的关联(随着年龄增长而产生的自我感知健康程度下降对男性的幸福感产生了明显的正向年龄效应)。

由此看来,我们认为死亡率差异可能并不是本研究中显示的、年龄和世代模式的原因。然而,如果很难找到对于这

一发现其他的原因，这对于男性从70岁到80岁之间幸福感的大幅上升是一个很有吸引力的解释。

在此，我对数据进行的非正式检验为过去几十年中美国男性和女性的年龄对幸福感影响的方向和大致特征提供了有力的证据。这些效应的大小是很难进行精确测量的，因为它们在截面数据中与世代效应相混淆，且在世代内趋势数据与时期效应相混淆。我们可以使用统计模型来获得准确的估计，但是估计将有很大的误差，而只有当它们与根据非正式的方法获得的大致模型是一致的时候，估计才是可信的。另外，根据精确估计获得的结果并不一定比利用有关效应的方向和大致模式的信息更加有用。

在对年龄和幸福感的全面研究中，对估计出来的效应提供理由和相关的证据自然是非常重要的。在此，我简单地指出，年龄对男性和女性幸福感具有不同模式的影响的一个可能的原因就是，男性和女性在同一年龄所处的婚姻状态可能不同。对任何年龄的男性与女性，已婚人士的幸福感都高于未婚人士。这种相关可能部分归因于选择结婚或者不结婚的人群的特征，但这当然在一定程度上是结婚对幸福感的正效应的结果。在较年轻的时候，女性比男性更倾向于结婚，但是在年纪较长时，男性比女性更倾向于结婚。根据1972年至2002年的美国综合调查数据，男女结婚人数比例的差异在19岁到76岁之间与年龄有着极大的线性相关（$r=0.95$，以每一年的年龄水平作为分析单位）。男女表明自己“非常幸福”的人数比例的差异同样与年龄有着较强的线性相关（$r=0.792$），而男性与女性在结婚比例和表明自己“非常幸福”比例的差异之间的相关系数为0.823。同样是用每

一年的年龄水平作为分析单位，且分析的样本年龄在 19 岁到 76 岁之间(为了使每个单元格中的样本数量大于 100)，对男女表明自己“非常幸福”比例的差异根据年龄进行回归的回归系数为 0.253($p < 0.1$，双尾检验)，但是在加入男女结婚人数比例的差异后，年龄的系数下降为 0.033(下降了 87%)。当控制年龄时，对男女幸福感的差异根据结婚比例的差异进行回归的回归系数为 0.263($p < 0.01$，双尾检验)。

另一种解决这个问题的方法是，对男性和女性分别根据年龄对幸福感进行回归，之后加入婚姻状况这一协变量，比较它对回归系数产生的影响。当对整个 1972 年至 2002 年的数据进行回归时，男性和女性样本中年龄的系数分别为 0.004和 0.001，它们的差别为 0.003($p < 0.001$，双尾检验)。在这个模型中加入了婚姻状况之后，男性样本中年龄的系数下降为 0.001，但女性样本中年龄的系数则几乎没有变化(因为女性的年龄与婚姻状况的线性相关几乎为 0)。

这些数据显示(尽管不能完全的证明)，男性与女性在婚姻状况上年龄的差别在很大程度上导致了男性与女性的年龄效应对幸福感影响的不同。

第3章

用世代分析理解变革

以下的世代分析仅限于人口的出生率、死亡率以及老龄化速度没有大的变化的同代。在这样相对封闭的人口中,从时间点 A 到时间点 B,群体特征的变化主要通过三种形式:(1)通过从初生和成熟的而增加的个体;(2)通过老龄化和死亡减少的人口;(3)在时间点 A 与时间点 B,除去由老龄化的影响,群体中个体特征的变化。前两者构成“世代演替”,而最后一种叫做“世代内的演变”。[3]

在《世代分析》的第一版中(Glenn, 1977),我认为将人群按照上述三种变化模式细分,对于分解研究人口是非常有帮助的,并且建议发展相关的研究方法。至少已经有一个研究者格伦·菲尔鲍,响应了我的提议并且设计出了几种将人口按照上述变化分类的方法(Firebaugh, 1989、1990、1992、1997)。

在这里,我不对菲尔鲍的方法进行具体解释,因为我现在相信,无论是它们还是任何近似的分解方法,对于理解改变都不是很有帮助(Rodgers, 1990)。这种分解研究仅对不具年龄效应的因变量有意义,即使在极为罕见的情况下,即当一个人高度自信地认为根本没有什么年龄效应时,这种分解的结果是否有效也在很大程度上取决于它所覆盖的时间范围。

用一下一个假设的案例便可以说明年龄效应是如何使

得分解人口变化变得毫无意义。我们假设在 25 岁到 64 岁的人中，每衰老 10 年便获得＋10 的因变量。在时间点 A 与时间点 B 之间，年龄分布不发生变化，则从时间点 A 到时间点 B 只有＋10 的时期效应，并且随着时间的推移，“世代演替”也没有任何作用。从另一方面来说，年龄效应对于年龄分布不变人群的人口变化不发生影响，而对“演变”却存在某种影响(除时期效应以外)。因此，在这个假想的例子中，演变对于解释人口变化已经绰绰有余。如果利用菲尔鲍设计的或任何相近的方法，即使事实上世代演替并没有什么效应，也会显示出世代演替对于人口变化显著的负面影响。

另外一个问题是，虽然世代演替在一个相对短的时期里对人群的变化几乎不起作用，可是如果时间足够长，则任何分解研究都会把人口变迁归因于世代演替。设想一下，同样以 25 岁到 64 岁的人群为例。40 年之后，没有一个初始成员还会留在这个群里，自然而然，在这里没有任何人口变化可以归因于世代内的变化。而如果我们一定程度地细化时间，比如，在 25 年到 35 年之后，这将会改变由世代演替对人口变化产生的效应以及对可见的人口变化的解释(尽管它不应该产生这样的效果)。

即便如此，如果能得到合适的数据，世代的视角对于理解社会、文化以及政治上的变迁总是有用的，例如，由于世代连续性的存在，世代分析能够识别持续下来的变化，即使这些变化的影响已经解释了很久。如果像某些理论以及大量证据表明的那样，通常时期效应对年轻人比对年长者更具有影响力，那么足够大的时期效应将会随之造成显著的人群内差别。

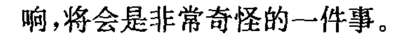
表 3.1 美国不同时期不同年龄人群对性行为态度的回归分析

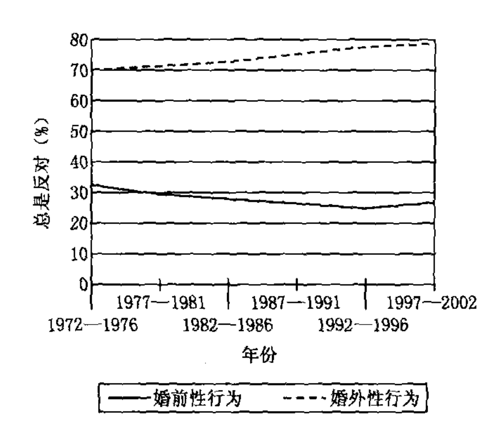
时期	婚前	婚外
1972—1976	−0.022***	−0.010***
1977—1981	−0.022***	−0.007***
1982—1986	−0.015***	−0.004***
1987—1991	−0.019***	−0.004***
1992—1996	−0.016***	0.001
1997—2002	−0.014***	0.000

注:*** 表示 $p<0.001$(双尾检验)。
资料来源:由 1972 年至 2002 年美国综合社会调查的数据估算得出。负系数表明,年长者比年轻人更为保守。

20 世纪 60 年代至 70 年代早期美国“性革命”的影响似乎就说明了这样一种情况。在 20 世纪 70 年代中期,即这场运动的大发展时代,社会普查中关于不同性行为是否“有错”的答案出现了相当大的“年龄差异”:在行为模式上,年轻人比年长者更宽容自由(表 3.1),尤其在婚前性行为这一问题上。这些差异本来也许能反映出年龄效应,因为人们随着年龄增长可能性态度会越来越保守。不过,在下文所列的证据中,这看上去仅仅是由不同年龄的人对性革命的刺激有不同反应而引起的。

20 世纪 70 年代中期的这种人群内部差异有助于我们理解一个似是而非的说法,即从 20 世纪 70 年代中期开始到 20 世纪末,人们对于婚前性行为与婚外性行为呈现出截然相反的态度(图 3.1),在所有的美国人中,人们对于婚前性行为的态度越来越宽容,而对待婚外性行为的态度则越来越严苛。这两种趋势在统计数据上都是非常显著的($p<0.001$ 双尾检验)。这两种趋势看似互相矛盾,因为它们彼此相关($p=0.301$)并且与其他变量有类似的相关性。因此,如果时期效应对这两种发生于过去几十年内且截然相反的趋向产生影

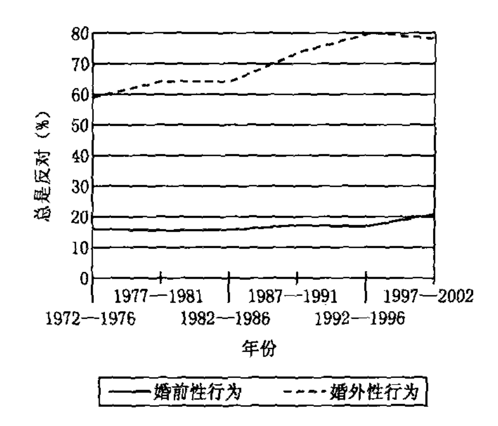

响，将会是非常奇怪的一件事。



资料来源：由1972年至2002年美国综合社会调查的数据估算得出。

图3.1 美国不同时期对婚前及婚外性行为表示“总是反对”的18岁以上人群的百分比

时期效应不太可能以彼此反向的形式出现。正如笔者上文所述，时期效应的趋势通常体现在年轻人群的特征中。18—29岁人群从1972年到2002年对婚前性行为和对婚外性行为的态度都越来越严苛（图3.2），这种趋向在人们对婚



资料来源：由1972年至2002年美国综合社会调查的数据估算得出。

图3.2 美国不同时期对婚前及婚外性行为表示“总是反对”的18岁至34岁人群的百分比

外性行为以及婚前性行为的态度中非常显著,而在前者中体现得更为明显($p < 0.05$ 对于婚前性行为; $p = 0.001$ 对于婚外性行为)。

综上所述,几乎可以确定地说,对整个人口而言,由于世代演替,人们对于婚前性行为的态度将会变得宽松,并且这种世代演替的效果强大到可以抵消时期效应所带来的反作用力(更加保守)。与之相反,人们对婚外性行为的态度趋向于保守,并且时期效应的作用也足够强以至于可以抵消世代演替在这方面的影响。这其中也许有一种趋向严格的年龄效应(年龄越大,态度越保守),也许没有,但是即使有,也会被时期效应或者世代演替的作用掩盖。

了解这些设定时间段内的人群内曲线总是有益的,1972 年至 2002 年这 30 年内的世代曲线在表 3.2 以及图 3.3、

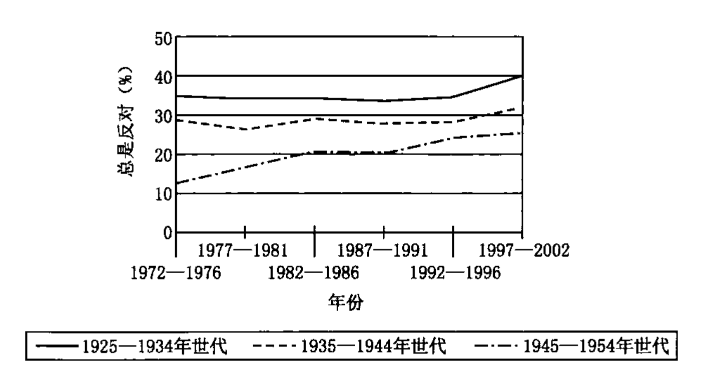
表 3.2 美国不同年代的出生人群对性行为态度宽严对比的年度(1972—2002 年)分析

世代(出生年份)	婚前	婚外
一般最小二乘法[a]		
1925—1934[c]	−0.002	−0.008***
1935—1944[d]	0.002	−0.008***
1945—1954[e]	−0.012***	−0.011***
logistic 回归[b]		
1925—1934[c]	0.004	0.029***
1935—1944[d]	0.006	0.034***
1945—1954[e]	0.030***	0.035***

注:a. 负系数显示了越来越严苛的态度趋向。
b. 人们的回应被分为"总是反对"对"其他态度",正系数显示了越来越严苛的态度趋向。
c. 本数据覆盖了年龄从 38—47 岁开始到 64—73 岁为止的世代。
d. 本数据覆盖了年龄从 28—37 岁开始到 54—63 岁为止的世代。
e. 本数据覆盖了年龄从 18—27 岁开始到 44—53 岁为止的世代。
*** 表示 $p < 0.001$(双尾检验)。

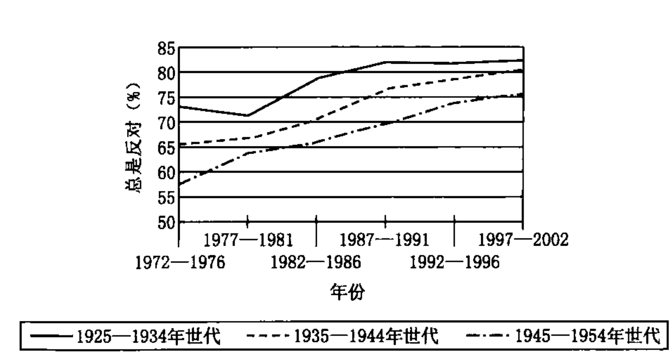
资料来源:由 1972—2002 年美国综合社会调查的数据估算得出。

图3.4中都有体现。我使用根据人们态度宽严差异的二分法(总是反对相比于其他态度)进行回归分析。并且采用了顺序度量作为这张表上半部分的最小平方回归分析的区间。



资料来源:由1972年至2002年美国综合社会调查的数据估算得出。

图3.3 美国不同时期及不同年份出生的人群对婚前性行为表示"总是反对"的百分比



资料来源:由1972年至2002年美国综合社会调查的数据估算得出。

图3.4 美国不同时期及不同年份出生的人群对婚外性行为表示"总是反对"的百分比

世代内婚前性行为态度的意义并不是很清楚。只有最年轻的群体才在统计中显示出显著的趋势,其他两个世代在1992年至1996年以及1997年至2002年的时间段里都趋向

于严格。我和哈丁以及詹克斯(Glenn, Harding & Jencks, 2003)一样,通过类似的数据来说明人们从刚成年到30岁时会因为年龄小而导致态度越来越保守,但在全部三个世代中,在此时期内这种类似的变化却也很有可能是由时期效应造成的(表3.3)。无论如何,时代内并没有出现越来越宽容的趋势,而且在年龄的成年人中也没有出现这一趋势,都显示出整个成年人口中越来越宽容的趋势完全是由于世代演替造成的,而非时期效应的结果。

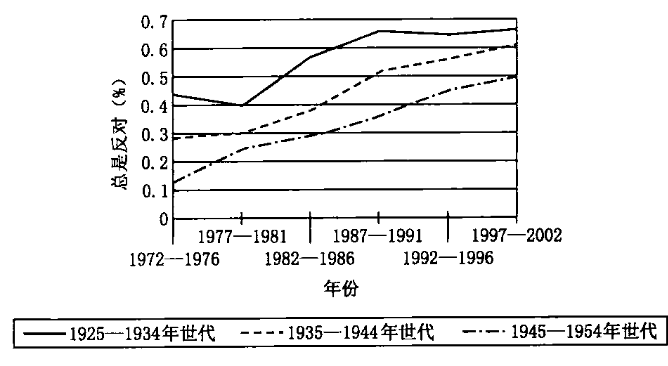
世代对婚外性行为的态度趋势的意义则更加清晰。三个世代都显示出强烈而明显的越来越严苛的态度趋向(表3.2及图3.4)。通过观察整个成年人口和年轻成年人中的类似趋势(图3.1),我们可以推测这样的变化很大程度上是由时期效应引起的,尽管年龄效应在这里不能被排除(随着年龄增长,态度趋向严苛)。在这里,趋向严格的时期效应必须足够强大,除了能够形成明显可见的趋势外,它还需要抵消由世代演替所引起的在该分析时期的早期所产生趋向宽松的反效应。在表3.1中,在“1972—1976”以及“1987—1991”的时间段内,年龄与采取严苛态度之间的正相关显示了这样趋于宽容的效应。

尽管如此,如表3.1所示,在20世纪90年代中期,年龄与婚外性行为之间的对应关系却消失了。这个变化也许是由于青年人转而比年长者采取更加严苛的态度所造成的,但也有可能是因为刚成年的新世代所采取的态度与年长者更为相似。综上所述,这些发现其实与“对于可能导致变化的刺激,年轻人群更容易受影响”这个假设是一致的,但图3.4中的趋向曲线却没有显示出三个世代的单调收敛性。此外,由

于天花板效应，最年长组比起年轻组，趋向变化要少一些，也就是说，由于该组成员的态度实际已经非常严苛，这使得一些测量工具无法测量出他们是否变得更严苛。

当对一个连续变量用二分法进行测量时，即将对性行为态度的测量分为“总是错的”和“其他”两种答案时，将百分比转化为 logit 形式是可以帮助修正“天花板”和“地板”效应的。这个转换包括将不平等百分比转换为发生比，并对发生比进行对数转换。基于连续变量的分布大体是正态的这一假设，相比于 50 左右的百分比差异，这个过程给接近 0 或 100 的百分比差异更大的权重。当然，这种“修正”的正确性取决于假设的正确程度。

与图 3.4 中表现的百分比对应的 logit 转换结果展现在图 3.5 中。



资料来源：由 1972 年至 2002 年美国综合社会调查的数据估算得出。

图 3.5 美国不同时期及不同年份出生的人群对婚外性行为表示“总是反对”的 logit 转换

在 1972 年至 1976 年间，该数据仍然体现出世代内差异的净减少，但它同时也体现出不同世代间的差异（1977 年至 1981 年及 1987 年至 1991 年间）。因此，通常意义上来说，这

样的数据过于“凌乱”以至于不能支持该假设。造成这种情况的原因是多种多样的，除了年龄效应被复杂化以外，还可能是因为年长者对于变化刺激的反应其实与年轻人相一致(只是不那么迅速)所造成的。

第4章

数据要求与数据可用性

第1节 | 数据要求

我们通常都在两个或更多时间点上来收集世代分析的数据,而且最有用的数据通常都覆盖几十年的时间跨度。很显然,世代分析不会按照教科书中的顺序来制定研究计划、收集数据和进行分析,因为它的数据形式和数据可用性都是由其他学者根据其他目的而收集的。换句话说,它们其实只是二手数据。反之,有用的数据通常都是从样本研究中直接获得的,尽管其中有些数据(比如人口普查资料)原本是针对所有人而非特定人群进行的。

世代分析的一个重要要求就是,所研究的人群要近似于"封闭",也就是说,除了出生和死亡人口以外,人口流动性要较小,或者当使用划分年龄的子样本的数据会通过老龄化而产生变动,比如使用25岁至64岁的人群作为研究对象。用其他方法所导致的人口的变动可能使人口年龄特性变化出现假象,或者掩盖已经发生的变化。因此,地区、州(美国的州)、社区以及类似的地区不适合进行世代分析研究,那些已婚人士、雇员或者有高度渗透性边界的群体也不适合。尽管没有什么人是完全封闭的,但单一民族国家的成年人却异乎寻常地适于进行世代分析,除非这个国家经历过大规模的移民。根据性别、种族、民族甚至宗教进行分组的人口也同样足够封闭。当然,根据

成年前的背景资料进行的人口分组在其成年后也是封闭的。

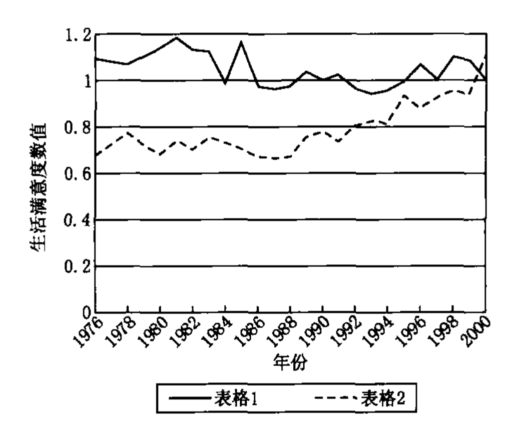
世代分析的第二个重要要求就是，在不同时间点所收集的数据必须具有可比性。比如调查问卷的数据，所有题干和可供回答的选项必须保持一致，除非有证据证明这些不一致不会导致反应类型的变化。问卷措词的微小不同可能导致反应类型大相径庭，所提供的选项顺序不同也会导致结果有很大的出入。比如，在一个关于生活满意度的调查问卷中，调查对象被要求按照从“完全满意”到“完全不满意”的几个数值范围中进行选择。满意度排序的最先选项可能导致调查对象的选项发生变化。数值范围的中间选项是用言语表示还是用数字表示也可能会导致结果的不一致。

即使调查问卷的题干的措词和选项都一样，结果还是有可能有差别。由不同的问卷调查组织收集的数据可能因为样本设计、访谈人士的训练和监管、编码程序（机构效应）而不同，通过不同调查方式（面对面、电话、自行管理）所得的数据，严格意义上并不具可比性（模式效应）。更重要的是，早前被问到的问题，尤其是刚问过的问题，会极大地影响调查对象的答案（问卷语境效应）。比如，关于总体生活满意度的问题是否出现在关于某个领域生活满意度问题之前，所导致的选择很可能很不一样。

很明显，用可以比较的数据追踪世代间的趋势，或用某一年龄层内连续的世代演替来追踪单一年龄层内的变动趋势。而不明显的是，即使可比较的数据也会给未发生的趋势制造已经发生的假象或者掩盖真实趋势。这一点在图 4.1 中已经详细举例说明了。该图显示，在关于高中女生对生活满意度的调查问卷中，由密歇根大学社会研究所进行的“监

控未来调查"中用两份不同的二次抽样样本所得到的选择结果存在组别趋势的差异。

在年度调研的表格 1 中,学生被要求使用包含 7 个数值的量表对其生活满意度进行评估,该数值范围将"完全满意"放在第一位,最后是"完全不满意"。数值范围的中间值是"中性",数值范围的其他点则仅仅使用了数字而不是数字和文字并用。在询问了关于生活的 11 个具体方面的满意度之后,包括从"你的教育经历"到"你是如何打发闲暇时间的",学生们被要求评价"生活整体满意程度"。在表格 2 中,调研人员先对二次抽样中独立抽出的学生们问了许多完全不相关的问题,然后询问他们:"最近你的生活整体满意度如何?"问题的选项先列出了"完全不满意",然后才过渡到"完全满意",并且这 7 个等级数值范围的中间值包含了数字和书面标签,比如"相当满意"或"还算满意"。在图 4.1 中,我给"完



资料来源:由 1976 年至 2000 年"监控未来调查"问卷得出。以 7 分为生活满意度标准值,完全满意则加 3 分,完全不满意则减 3 分,表格 2 用文字标注的形式来标示所有 7 个问题。

图 4.1 美国高中女生年度生活满意度值

全满意”标上+3分，“完全不满意”标−3分，中间值标0分。每年每张表格的平均反应值都被绘制在数据中。

尽管表格1和表格2的问题在名义上是一样的，但不仅平均反应值不同，而且几十年来的发展趋势也朝相反的方向进行。当使用问卷总数进行分析时，表格1中的问题所得出的平均值与年份的相关系数是−0.458，是具有统计意义的（$p<0.05$，双尾检验），而表格2的平均值与年份的相关系数是+0.793（$p<0.01$，双尾检验）。很明显，这两个相关系数都不能准确指明美国高中高年级女学生生活满意度的最新发展趋势。

出现这两种相反趋势变动的原因尚不知晓，因为这两份关于生活满意度的调查问卷有3个地方不同。第一也许是最重要的一处，表格1中关于生活整体满意度评价的问题是在问了生活的11个相关方面满意度的问题之后提到的。第二，表格1的该选项是按照“完全满意”在先，过渡到“完全不满意”的。第三，表格1仅在局部范围内用数字来标示7个选项中的4个，而表格2是用文字形式标示所有7个选项。

这种对两个问题应答的不同趋势很有可能（虽然不确定）是由提问顺序效应引起的，因为对于表格1内一部分特定问题的回答显示出向下的趋势。也就是说，前述问题可能引起了人们对生活中满意度下降的思考。回答问题中的一种趋向可能会产生一种延续该趋向的表象或是假象，但如果在后续阶段受到提问顺序效果的影响，则有可能产生变化。

从这个例子中我们了解到，如果对因变量的测量很容易受到问题顺序效应的影响，则应该在世代分析中尽量避免这类测量。如果使用了这种方法，得出的分析结果应当标注明

显的警示。

由于样本设计的变化,由同样的机构按照同样的形式(例如面对面调查)在不同时期收集的数据也许是不具有可比性的。正如我在下文中将指出的,在 20 世纪 30 年代、40 年代以及 50 年代所进行的那些民意调查对于分析趋向以及世代分析而言,都是有用的资源,但不幸的是,早期民意调查所收集的那些数据与新近的数据不具有可比性,即使是由同一个机构收集的数据。早期调查是以控制配额的方式来设计样本,并按照由年龄、性别、地区对不同人群进行配额。例如,一个调查可能需要 10 个男性受访者、10 个女性受访者、6 个 35 岁以下的受访者、8 个 35 岁至 39 岁的受访者以及 6 个 60 岁及以上的受访者。考虑到这些限制,调研者就已经确定了寻找目标受访者的大体方向,而这个致命的缺陷会导致严重的偏误。到 20 世纪 50 年代后期,尽管全概率样本在美国人口普查局以及类似的学术机构中(如密歇根大学的调查研究中心)还没有被使用,但大多数民意调查都已经采用了更为精密的抽样设计方法。最典型的例子是,一种类似于全概率抽样的分阶段地区抽样过程被用于街区层级的抽样,但是在选定受访者的最后阶段,配额选择仍然是重要的方法。这种形式的样本被称做"改良概率抽样"或"配额概率抽样",与已经确定了受访者选择方式的配额控制抽样有重大的区别,并且访谈中不存在选择受访者的问题。严格地说,用这样的方式取得的数据与配额数据抽样和全概率抽样都不相容,但很显然,它们与后者可能更接近一致。

美国民调抽样方式最重要的改变出现在美国盖洛普民意测验机构,这个机构早年叫做"美国民意调查研究所"。盖

洛普第一个伟大的成就是预测了1936年总统选举的结果，在此后的20年间，这个机构一直拥有敏锐的政治眼光并主要关注总统选举。盖洛普早期的样本自然也是为代表选民而不是全体成年人口设计的。在这样的样本里，那些选举权较低的选民明显代表名额不足，像妇女、黑人以及没有受过高等教育的人。自20世纪50年代早期开始，盖洛普的样本开始改良，这个过程直至1958年。此次改良计划使选择其他应答的受访者的百分比上升8到10个百分点。

盖洛普所收集的数据曾经涵盖了大量趋向研究和世代分析可用的数据，我曾经建议采用多重标准化程序以使盖洛普早期和晚期收集的数据有可比性。这些数据有时会很有用，但随着不同时期、不同主题（社会学家们感兴趣的）的高水平学术调研所收集的数据越来越多，盖洛普的数据已经不像以前那么有吸引力了。因此，笔者在此不再讨论采取技术手段使盖洛普早期和后期收集的数据相匹配的问题。大多数世代分析所用的数据也不存在这个问题。

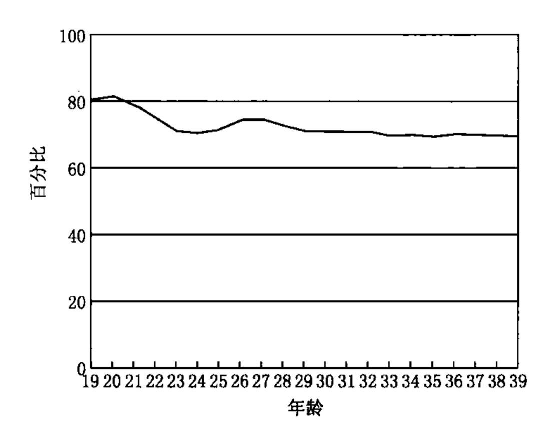
还有一些问题不像盖洛普早—晚期数据匹配这么严重，它们是由主流民意调查机构将其调研方式从面对面改为电话采访而引起的。这些问题在20世纪70年代末到80年代初尤为典型，虽然它们并未严重到不足以影响对于同一机构所收集的早期数据和晚期数据的运用，但研究人员必须了解什么时候该用电话调研方式，并且要寻找那个时期对一些重复问题的应答的突然转变。

在美国的大多数民意调查中，年轻人并没有像年长者那样充分被代表，由此引发的一个可比较性问题经常被人们忽视（包括我）。几乎所有的主流民意调查样本都只研究没有

被制度化的人群，而“成群驻扎”的人们并没有被列入研究对象，它包括监狱、医院、军营以及学校宿舍等地方，除医院以外，此类地方都有很多年轻人。这种由于上述群组年轻人代表名额不足所引起的偏差，在某种程度上是可抵消的，因为从贫困的学生到中高阶层的学生都被排除在样本之外。但是，对某些变量来说，这种净偏差是不可忽略的。

研究世代分析的学者应当注意这种潜在的变化并注意这种存在于成年期早期及中晚期之间的变量的变化。不过，这个问题对大多数世代分析来说，似乎并不重要。老年人与年轻人世代是否可以进行比较，则可以通过观察随着年龄的增长，世代内背景特征是否相应变化来进行评估。任何群体内显示的变化如果不是抽样错误的结果，就一定是样本间没有可比性的结果。基于这种估计，运用第 2 章所提到的关于幸福感分析的程序以及对同样年龄范围内(不是同样时间范围内)的个体进行以 1 年为单位的跟踪，而非整个相同的时期的追踪。我运用这个步骤，分析了几个未成年背景特征，没有发现这种人群内表现出性观念的变化(随着年龄增长)。在图 4.2 中，我们可以看到这样一个例子：自 1955 年至 1964 年，在美国出生的世代从 30 岁开始就被追踪调研，并且我们假设这个人群的个体中，只有一位接受过高中教育的母亲。在此背景特征下，表示性观念变化的百分比自个体接近 20 岁到 25 岁左右的阶段出现了一定程度的下降，这大概是因为大学生回归到了样本人口的范围中，并且在 30 多岁时依然保持该状态。这种已经显示出来的细微变化对于世代分析的结果不会产生很大影响。尽管如此，那些对于年轻人年龄效应没什么特别兴趣的学者还是倾向于用 25 岁而不是 18

岁作为它们分析的年龄起点，以降低由年龄相关样本的不可比性而引发的偏倚的概率。



资料来源：由1972年至2002年美国综合社会调查的数据估算得出。

图4.2　美国1955年至1964年出生的世代中受访者母亲接受过高中教育的百分比（根据年龄，每3年计算一次均值）

第2节 | 数据可用性

本书中,我在例子中引用的数据大多来源于美国综合社会调查(在1972年至2002年间每年或每两年进行一次)。美国符合社会调查中有很多重复性的问题我并未使用,有一些问题是从1972年到2002年全部包括的,有些是从1973年开始的,还有的至覆盖了很短的一个时期。在这些问题中,有的问题在每个调查中都会出现,而有的则以一种螺旋式的方式呈现,例如在两个连续的调查中被提出,结果该问题在下一次社会调查中不出现,然后又出现在两个新的连续的调查中。近年来,社会普查数据每两年进行一次,其中的样本自然是早些年样本大小的两倍,并且这些数据涵盖了很多只有一部分受访者会被问到的问题。

美国综合社会调查中的一些更深入的问题涉及对社会生活各方面的满意程度、对当前形势的信心、思想意识取向、政党认同感、投票参与总统竞选的情况、工作满意度、工作态度、疏远感、志愿者组织成员资格、对孩子成就的期望、对堕胎的态度、词汇量、宗教偏好、宗教信仰、民族态度以及吸烟、饮酒状况等。和大多数面对面调查以及电话调查一样,美国综合社会调查的抽样对象是家庭而非个人,因此,这样的样本必须根据家庭中个体数量被重新加权,以确定它能够代表

未被制度化的成年人群体(没有被重新加权的数据过高代表了只有一个成年人的家庭中成年者的数量,过低代表了多个成年者家庭中成年人的数量)。为了保证研究中显示的受访者与实际受访者数量一致,我们使用一个部分加权而不是原始的家庭中成年者数量。这个加权平均值由分割家庭中的成年者的数量除以美国综合社会调查中得到的家庭平均承认数量得出,其值为1.94(1972年至2002年间)。这个权数估算可以在每一个调查中进行,但它通常只在由所有调查数据构成的汇总文件中才能进行。在1982年和1986年的调查中包括了一个权数来处理黑人的过度抽样,但随着美国综合社会调查的文件越来越大,黑人的过度抽样变成相对非常小的一个部分,对此进行加权对结果几乎没有影响。纯粹的统计学理论研究者可以给每一次调查相等的权数,但这么做其实是拘泥于细节而没什么应用价值的,并且它对整个研究结果的影响可以忽略不计。

美国综合社会调查的数据可以通过多种来源获得,从洛普公众意见研究中心的光盘到社会及政治研究校际联合会的网站上都可以获得,还可以在洛普中心买到编码簿。美国选举研究(ANES)比美国综合社会调查提供了更多的重复提出的问题,包括好几个系列,比如是否缺乏政党认同感等等,这个问题从1948年第一次ANES进行以来,就一直包含在调查之内。其他几个"系列"从1952年开始,并且这些问题所覆盖的时间长度比美国综合社会调查中覆盖时间跨度最长的问题还要长。其中一些问题是每两年一问,而有些是每四年一问。这些问题大多是从政治视角出发,为政治参与度、政治体制的选择以及党派偏见提供了数据,但是社会学

者、社会心理学者、经济学者，还有政治家都可以从中获得有用的数据。例如，ANES问了许多关于政策方面的问题，涉及医疗保障、民权行动、堕胎以及军事、经济等方面。

ANES有自己的网站，通过选择所需信息表格的引导，你可以在上面下载你需要的数据，并且加权信息也是给出的。在进行世代分析或者其他覆盖时间范围较久的社会分析以前，你必须好好研究这些复杂的信息。有些加权信息是从1994年才开始提供的，并且如果你要使用早年的数据，你就会发现这些不同时期的数据完全不具有可比性。

除美国之外，一些其他国家也在实践类似于美国综合社会调查和ANES的信息采集工程。我们可以从社会及政治研究校际联合会的网站上找到其中的一些数据，但有一些只能在美国以外的档案馆或网站上找到。

一个能给世代分析带来巨大好处并且已经收集了很多有用信息的数据收集项目是"世界价值调查"，一个关于社会文化与政治变迁的世界性调查。第一次世界价值调查从1981年至1984年，其后还有1990年至1993年、1995年至1998年以及2000年至2001年3次，每次调查都包括80个左右的国家，且每个国家包括超过1000名受访者。有20余个国家参加了所有4次调查并且提供了重复并具有代表性的数据，我们可以用这部分数据来进行世代分析。这项调查提出了一大批多种多样的问题，如环境问题、社会规范问题、对婚姻及家庭的态度、性观念问题以及政治态度等等。前三次世界价值调查的数据可以从社会及政治研究校际联合会的网站上找到，并且在本书即将完成的时候，第四次调查的数据将被提供给大众。关于这项调查的最新信息可以在世

界价值调查网站上找到。

尽管有其局限性,但商业民调数据对于世代分析,仍是一种十分有用但却未被充分使用的数据,它所提供的数据周期很短而不像一般社会普查那样是 1 年至 4 年(在此期间有很多重复的学术研究)。在美国,很难想象有什么主题不是以民意调查形式来进行的,并且有很多好的问题被重复发问了数年甚至几十年。民意调查主要由两个机构负责:康涅狄格大学的洛普公共意见研究中心以及北卡罗来纳大学教堂山分校中由奥德姆学院负责的路易斯·哈里斯数据中心。洛普中心是世界上历史最悠久也是最大的民意调查中心,它保有自 1935 年以来由民意调查机构所收集的绝大部分非专利数据。人们可以通过洛普中心的网站来搜寻数据,但数据仅对中心成员以及缴纳使用费者开放。另外一个主要的民意调查机构——路易斯·哈里斯中心——所收集的大部分数据,仅对奥德姆学院可用,这也使得后者保存了自 1958 年开始的超过 1200 个由哈里斯中心所做的民意调查的数据。人们同样可以在奥德姆学院的网站上搜寻数据,并且某些数据可以在线获取。其他各工业化国家至少拥有一家保存了丰富民意调查资料的机构。

不幸的是,在某些早期民意调查中,年龄至分了较大的类别,这使得它们对于世代分析的价值非常有限。要为世代分析寻找合适的民意调查数据,在消耗大把的时间和进行数据研究之前,人们还应当看看年龄是不是精确测量的。

第5章

未来展望

世代分析的前景是光明的,因为有大量的可用数据能够支撑这种研究继续前行,并且,现在其实已经有大量的数据可以帮助我们理解年龄、社会以及文化的变化。尽管不会有一种绝对理想的方法被应用于"年龄—时期—世代"难题,但仍然会有方法论上的不断改良,从而推动世代分析研究的发展。

采用精密的统计方法而不是本文所涉及的简单方法会对世代分析产生积极的作用,尽管这些方法仍不能解决"识别"问题。我们应当承认,这些努力对于其他社会研究和世代分析来说,都有其意义,但显然专门为世代分析而设计的研究方法一定会更有效。

不幸的是,在短时期内,这些其他的研究方法对世代分析的意义是有限的,因为分析中既要用到过去的数据,又要用到新近的研究结果。世代分析依然会受到"所使用方法不够先进"这样的批评。尽管如此,世代分析自身的高敏感性对解决其方法论问题仍然会有所帮助,其中一个例子就是本书第一版发行之后,大家对于问题顺序效应有更清晰的认识。

目前为止,几乎所有的世代分析都会用面对面调查这种

形式以获得数据，但几乎可以断定的是，这样的问卷调查以后会更依赖于其他方式以获得数据，例如通过电话调查来获得数据。这些以其他方式所获得的数据会引起一些问题，而笔者在专著中并未提及或仅仅是一笔带过。例如，与面对面调查相比，与年龄相关的抽样的不可比性在电话调查中可能是不同的。如上文所述，不可比性是由于在同一个研究中数据是由不同管理问卷的方法而导致的，并且会产生一些至少在不同时间利用电话调查一样严重的问题。为了进行电话调查，应答机与来电者身份确认制度的广泛运用可能会引起新的偏误，并且目前还没人知道越来越多的家庭成员仅仅依赖于手提电话而对调查结果产生的影响。

世代分析所面对的逻辑难题始终一样，然而它所涉及的技术手段却一定会改变，而且会在未来的几年里迅速改变。

注释

[1] 梅森和其他人没有指出在一个相对独立的可变年龄的范围内,对于一个线性变量,可能存在一个三变量解释。

[2] 人们可能会奇怪,为什么像群内成员父母离婚比率这样的数据也被用于个人层面的研究。这样做是基于一种假设的关联效应:父母没有离婚的儿童与青少年会受到父母离婚的同辈的影响,当然还有年龄关联效应以及时期关联效应。

[3] 年龄效应有时会被错误地列为可变因素之一。事实上,只有当群内年龄分布改变时,年龄效应才是相对可变的,这种情况在出生组(适合进行出生组分析)中随着队列演替而发生。

参考文献

Abramson, P. R. , & Engleghart, R. (1995). *Value change in global perspective*. Ann Arbor: University of Michigan Press.

Alwin, D. F. (1991). "Family of origin and cohort differences in verbal ability." *American Sociological Review*, *56*, 625—638.

Alwin, D. F. , Cohen, R. L. , & Newcomb, T. M. (1991). *Political attitudes over the life-span: The Bennington women after fifty years*. Madison: University of Wisconsin Press.

Bachman, J. G. , Johnston, L. D. , & O'Malley, P. M. (1976—2000). *Monitoring the future: Questionnaire responses from nation's high school seniors*. Ann Arbor: University of Michigan Institute for Social Research. Twenty-five volumes(Ordering of compilers' name varies.)

Blalock, H. M. , Jr. (1966). "The identification problem and theory building." *American Sociological Review*, *31*, 52—61.

Blalock, H. M. , Jr. (1967). "Status inconsistency, social mobility, status integration and structual effects." *American Sociological Review*, *32*, 790—801.

Converse, P. E. (1976). *The dynamics of party support: Cohort analyzing party identification*. Beverly Hills, CA: Sage.

Davis, J. S. , Smith, T. W. , & Marsden, P. V. (2002). *General social survey, 1972—2002* (machine readable data file). Chicago: National Opinion Research Center.

Firebaugh, G. (1989). "Methods for estimating cohort replacement effects." In C. C. Clogg (Ed.), *Sociological methodology 1989* (pp. 243—262). Oxford, UK: Basil Blackwell.

Firebaugh, G. (1990). "Replacement effects, cohort and otherwise: Response to Rodgers." In C. C. Clogg(Ed.), *Sociological methodology 1990*(pp. 439—446). Oxford, UK: Basil Blackwell.

Firebaugh, G. (1992). "Where does social change come from? Estimation the relative contributions of individual change and population turnover." *Population Research and Policy Review*, *11*, 1—20.

Firebaugh, G. (1997). *Analyzing repeated survey* (Sage University Paper series on Quantitative Application in the Social Science, 07—115). Thousand Oaks, CA: Sage.

Glenn, N. D. (1974). "Aging and conservatism." *Annals of the American Academy of Political and Social Science*, *175*, 176—186.

Glenn, N. D. (1975). "Trend studies with available survey data: Opportunities and pitfalls. In Social Science Research Council Center for Social Indicators." *Survey data for trend analysis: An index to repeated questions in U. S. national surveys held by the Roper Public Opinion Research Center* (pp. 6—48). Williamson, MA: Roper Center.

Glenn, N. D. (1976). "Cohort analysts' futile quest: Statistical attempts to separate age, period, and cohort effects." *American Sociological Review*, *41*, 900—904.

Glenn, N. D. (1977). *Cohort analysis* (Sage university Paper series on Quantitative Applications in the Social Sciences, 07—05). Beverly Hills, CA: Sage.

Glenn, N. D. (1980). "Values, attitudes, and beliefs." In O. G. Brim, Jr., & J. Kagan (Eds.). *Constancy and change in human development* (pp. 596—640). Cambridge, MA: Harvard University Press.

Glenn, N. D. (1989). "A caution about mechanical solutions to the identification problem in cohort analysis: Comment on Sasaki and Suzuki." *American Journal of Sociology*, *95*, 754—761.

Glenn, N. D. (1994). "Television watching, newspaper reading, and cohort differences in verbal ability." *Sociology of Education*, *67*, 216—230.

Glenn, N. D. (1998). "The course of marital success and failure in five American ten-year marriage cohorts." *Journal of Marriage and Family*, *60*, 569—576.

Glenn, N. D. (2003). "Distinguishing age, period, and cohort effects." In J. T. Mortimer & M. J. Shanahan(Eds.), *Handbook of the life course* (pp. 465—476). New York: Kluwer Academic/Plenum.

Harding, D. J., & Jencks, C. (2003). "Changing attitudes towards premarital sex: Cohort, period, and aging effects." *Public Opinion Quarterly*, *67*, 211—226.

Kiecolt, K. J., & Nathan, L. E. (1985). *Secondary analysis of survey data* (Sage University Paper series on Quantitative Applications in the Social Science, 07—053). Beverly Hills, CA: Sage.

Mason, K. O., Mason, W. M., Winsborgough, H. H., & Poole, K. W. (1973). "Some methodological issues in the cohort analysis of archival data." *American Sociological Review*, *38*, 242—258.

Nakamura, T. (1982). "A Bayesian cohort model standard cohort table analysis." *Proceedings of the Institutes of Statistical Mathematics*, *29*, 77—97(in Japanese).

Nakamura, T. (1986). "Bayesian cohort models for general cohort table analysis." *Annals of the Institute of Statistical Mathematics*, *38*(Part B), 353—370.

O'Brien, R. M. (1989). "Relative cohort size and age-specific crime rates: An age-period-relative-cohort-size model." *Criminology*, *27*, 57—78.

O'Brien, R. M. (2000). "Age-period-cohort-characteristic models." *Social Science Research*, *29*, 123—139.

O'Brien, R. M., Stockard, J., & Isaacson, L. (1990). "The enduring effects of cohort characteristics on age-specific homicide rates: 1960—1995." *American Journal of Sociological*, *104*, 1061—1095.

Palmore. E., & Jeffers, F. C. (1971). *Prediction of life span*. Lexington, MA: D. C. Heath.

Rodgers, W. (1990). "Interpreting the components of time trends." In C. C. Clogg (Ed.), *Sociological methodology 1990* (pp. 421—438). Oxford, UK: Basil Blackwell.

Sasaki, M., & Suzuki, T. (1987). "Changes in religious commitment in the United States, Holland, and Japan." *American Journal of Sociology*, *92*, 1055—1076.

Sasaki, M., & Suzuki, T. (1989). "A caution about the data to be used for cohort analysis: Reply to Glenn." *American Journal of Sociology*, *95*, 761—765.

U. S. Census Bureau. (1969). *Statistical abstract of the United States: 1969*, Washington, DC: U. S. Government Printing Office.

U. S. Census Bureau. (1979). *Statistical abstract of the United States: 1979*, Washington, DC: U. S. Government Printing Office.

U. S. Census Bureau. (1990). *Statistical abstract of the United States: 1990*, Washington, DC: U. S. Government Printing Office.

U. S. Census Bureau. (1999). *Statistical abstract of the United States: 1999*, Washington, DC: U. S. Government Printing Office.

译名对照表

additive model	叠加模型
age cohort	年龄世代
age effect	年龄效应
Age-Period-Cohort model(APC)	年龄—时期—世代模型
Age-Period-Cohort-Characteristic(APCC)	年龄—时期—世代—特征
birth cohort	出生世代
cohort	世代
cohort diagonals	世代对角线
cohort succession	世代演替
compositional effect	组成效应
control group	控制组
cross-sectional study	截面研究
identification problem	识别问题
intracohort change	世代内的演变
panel conditioning effect	条件习惯效应
panel study	固定样本跟踪研究
practice effect	练习效应
quasi experiment	准实验
question-order effects	问题顺序效应
quota control sampling	控制配额抽样
randomized experiment	随机实验
side information	附加信息
small birth cohort	小型生育世代
standard cohort table	标准世代表

Cohort Analysis (2nd edition)

上海市版权局著作权合同登记号:图字 09-2009-549